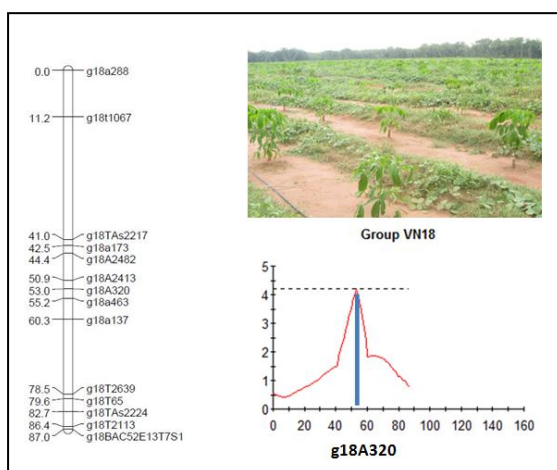


Master 2 - Recherche  
Biologie des Plantes et Micro-organismes, Biotechnologie, Bioprocédés  
Parcours Biologie des Plantes

2012-2013

**Optimisation de la cartographie génétique de familles F1 d'hévéa  
(*Hevea brasiliensis*) pour la détection de QTLs et le  
développement de la sélection assistée par marqueurs**



**Dinh Minh TRAN (8 juin 2013)**

Stage effectué au Cirad, Département Bios, Umr AGAP (2 janvier - 8 juin 2013).

Encadrement Cirad : V. Le Guen

Mots-clés : *Hevea brasiliensis* – Cartographie génétique – Microsatellites (SSR) – Electrophorèse capillaire – Détection de QTLs – Sélection Assistée par Marqueurs (SAM).

Keys words : *Hevea brasiliensis* – Genetic mapping – Microsatellites (SSR) – Capillary electrophoresis – Detection of QTLs – Markers Assisted Selection (MAS).

Résumé : 209 caractères

Abstract : 208 caractères

Texte : 7490 mots

## Résumé

Ce stage avait pour but de créer les référentiels génétiques nécessaires à la sélection assistée par marqueurs (SAM) chez deux familles biparentales d'hévéa et d'initier la détection de QTLs pour un caractère de croissance. Le génotypage simultané de 380 individus provenant de 2 croisements différents (PB260 x RRIM600 et PB260 x RRIC100) avec 403 marqueurs microsatellites a permis la réalisation de 2 cartes génétiques saturées, la première portant sur 190 individus avec 304 marqueurs (2004 cM), et la seconde sur 380 individus (190 ayant déjà été partiellement génotypés lors d'une précédente étude) et 354 marqueurs (1916 cM). Les trois cartes parentales ont également été générées. Dans tous les cas on observe une bonne colinéarité entre les marqueurs et un positionnement sur la carte conforme aux résultats obtenus antérieurement sur d'autres familles. Le phénotypage de ces familles expérimentées en champ en Afrique de l'ouest a permis la détection d'un QTL lié au diamètre au collet dans le jeune âge pour la famille PB260 x RRIM600, mais aucun QTL pour l'autre famille. Les données de cartographie seront exploitées dans les trois prochaines années pour effectuer la SAM sur ces deux croisements sur des critères de croissance, production précoce, et résistance à une maladie de feuilles causée par le champignon *Corynespora cassiicola*.

## Abstract

The aim of this study was to generate the genetic data necessary for the implementation of marker-assisted-selection (MAS) onto two biparental rubber families and to initiate the QTL detection for growth characters. Simultaneous genotyping of 380 individuals from two crosses (PB260 x RRIM600 and PB260 x RRIC100) with 403 microsatellite markers allowed to carry out two saturated genetic maps, the first one on 190 individual genotypes with 304 markers (2004 cM), and the second on 380 individuals (thanks to data on 190 individuals already partially available from a previous study) and 354 markers (1916 cM). The three corresponding parental maps were also established. In all cases a good collinearity was observed between markers, and a positioning on the map in accordance with previous results obtained in other families. The phenotyping of these families planted in field experiments in West Africa allowed the detection of a QTL for stem diameter at young stage in the PB260 x RRIM600 family, but no QTL was detected in the other family. These mapping data will be used during the next three years for developing a SAM strategy onto these two crosses on characters of growth and precocious production, as well as resistance to the foliar disease due to the fungus *Corynespora cassiicola*.

## **Remerciements**

*En préambule à ce mémoire, je souhaite adresser mes remerciements les plus sincères aux membres de l'équipe « hévéa » du Cirad, Marc SEGUIN (responsable de l'équipe EGV: Evaluation, Gestion et Valorisation des ressources génétiques), mon encadrant principal Vincent LE GUYEN, et André CLEMENT-DEMANGE, au sein de l'UMR AGAP, ainsi que mes collègues vietnamiens à l'Institut de Recherche sur le Caoutchouc du Vietnam (RRIV). Leur collaboration et leurs échanges scientifiques m'ont donné une opportunité d'étude sur l'hévéa, ainsi que le financement nécessaire de ce stage de Master 2 qui a été assuré par l'Institut Français du Caoutchouc dans le cadre du projet IFC-Création Variétale.*

*Je tiens à remercier sincèrement Ronan, Hélène, Caroline et les autres personnes qui animent le Grand Plateau Régional de Génotypage de la Région Languedoc-Roussillon au Cirad.*

*Enfin, j'adresse mes remerciements à mes enseignants du Master BPMBB à l'UM2.*

M. Dinh Minh TRAN, auteur de ce rapport de Master 2, est chercheur au sein de la division « Amélioration génétique de l'hévéa » du RRIV (Rubber Research Institute of Vietnam). Le sujet de ce rapport fait partie d'un partenariat de recherche entre le RRIV, le Cirad, et l'Institut Français du Caoutchouc.

## Table des matières

<b>Résumé .....</b>	<b>i</b>
<b>Abstract .....</b>	<b>i</b>
<b>Remerciements.....</b>	<b>ii</b>
<b>Table des matières .....</b>	<b>iii</b>
<b>Définition des sigles .....</b>	<b>iv</b>
<b>Tableaux .....</b>	<b>v</b>
<b>Figures.....</b>	<b>v</b>
<b>1. Introduction .....</b>	<b>1</b>
1.1. L'hévéa et son amélioration génétique .....	1
1.2. L'utilisation de marqueurs génétiques chez l'hévéa.....	2
1.3. L'évolution des techniques de génotypage par microsatellites .....	2
1.4. La cartographie génétique .....	3
1.5. La détection de QTLs .....	4
<b>2. Résultats.....</b>	<b>5</b>
2.1. Rappel de résultats antérieurs et synthèse des résultats nouvellement acquis.....	5
2.2. Extraction d'ADN.....	5
2.3. Test de légitimité des descendants.....	5
2.4. Génotypage des individus des deux descendance.....	5
2.5. Cartographie génétique.....	6
2.5.1. Cartographie de la famille PBRIM .....	6
2.5.2. Cartographie de la famille PBRIC .....	7
2.5.3. Cartographie des 3 parents.....	7
2.6. Détection de QTLs.....	14
<b>3. Discussion.....</b>	<b>16</b>
3.1. Longueur des cartes .....	16
3.2. Alignements des marqueurs sur les différentes cartes .....	16
3.3. Variabilité des longueurs des segments entre deux marqueurs consécutifs .....	16
3.4. Détection du QTL g18-53 pour le caractère D9 sur la famille PBRIM.....	17
<b>4. Matériel et Méthodes .....</b>	<b>19</b>
4.1. Matériel végétal .....	19
4.2. Génotypage et cartographie .....	19
4.3. Dispositif expérimental de phénotypage en champ .....	19
4.4. Détection de QTLs.....	20
<b>5. Références .....</b>	<b>22</b>
<b>6. Annexes .....</b>	<b>25</b>
Annexe 1. Alignement des marqueurs sur les 18 groupes de liaison des cartes. ....	25
Annexe 2. Détection de QTL de D4 pour la famille PBRIM (Côte d'Ivoire).....	34
Annexe 3. Détection de QTL de D9 pour la famille PBRIM (Côte d'Ivoire).....	35
Annexe 4. Détection de QTL de D7 pour la famille PBRIC (Nigéria).....	36

## Définition des sigles

ADN	Acide DésoxyriboNucléique
AFLP	Amplified Fragment Length Polymorphism
BLUP	Best Linear Unbiased Prediction
CIRAD	Centre de Coopération Internationale en Recherches Agronomiques pour le Développement
CP	« Cross-Pollinated » (plantes allogames)
DArT	Diversity Arrays Technology
EM	Expectation-Maximisation
EST-SSRs	Expressed Sequence Tags-Simple Sequence Repeats
JoinMap 4.0	Logiciel de cartographie génétique, version 4
LOD	Logarithm of the ODds ratio
P <sup>33</sup>	Phosphore radioactif
PB	Prang Besar
PBRIC	Croisement PB260 x RRIC100
PBRIM	Croisement PB260 x RRIC600
PCR	Polymerase Chain Reaction
QTL	Quantitative Trait Locus
RAPD	Random-Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphism
RRIC	Rubber Research Institute of Ceylon
RRIM	Rubber Research Institute of Malaysia
RRIV	Rubber Research Institute of Vietnam
SAM	Sélection Assistée par Marqueurs
SNP	Single Nucleotide Polymorphism
SSH	Suppressive Subtractive Hybridization
SSRs	Simple Sequence Repeats

## Tableaux

Tableau 1. Répartition des effectifs de s 403 marqueurs SSRs par type de ségrégation, par famille et par parent.

Tableau 2. Répartition des 403 marqueurs et longueurs en cM des groupes de liaison sur les différentes cartes des deux familles.

Tableau 3. Répartition des marqueurs et longueur de groupe de liaison sur trois parents.

## Figures

Figure 1. Carte génétique de la famille PB260 x RRIM600

Figure 2. Carte génétique de la famille PB260 x RRIC100

Figure 3. Variations des distances génétiques des groupes de liaison des 3 parents (cM).

Figure 4. Alignement des marqueurs du groupe g10 entre les cartes des 2 familles et des 3 parents.

Figure 5. Détection de QTL de D9 (diamètre à 9 mois) sur le groupe 18 de la famille PBRIM

Figure 6. Variabilité des longueurs des segments compris entre deux marqueurs consécutifs.

Figure 7. Disposition de deux familles avec 190 individus + deux parents pour chaque famille sur une plaque 384

**Optimisation de la cartographie génétique de familles F1 d'hévéa  
(*Hevea brasiliensis*) pour la détection de QTLs et le  
développement de la sélection assistée par marqueurs**

**TRAN Dinh Minh**

## 1. Introduction

L'hévéa est une plante cultivée dont l'amélioration variétale est basée sur des observations phénotypiques et l'utilisation classique de la génétique quantitative. Les études de génétique moléculaire de diverses équipes au cours de la dernière décennie ont permis de générer des marqueurs génétiques en grand nombre qui pourraient être utilement mis à contribution pour accroître la précision et l'efficacité des programmes d'amélioration variétale. L'objectif de ce stage de Master 2<sup>ème</sup> année était d'explorer cette voie en réalisant simultanément la cartographie génétique de deux familles dans le but d'identifier des QTLs liés à des caractères d'intérêt pouvant être utilisés en sélection assistée par marqueurs (SAM). Ce travail de cartographie génétique a été complété par l'identification effective de QTLs liés à la vigueur sur une des deux familles qui a été plantée en dispositif expérimental. Ce stage fait suite à un travail réalisé précédemment (Tran, 2012) et qui a permis de définir les conditions de génotypage à fort rendement de l'hévéa en utilisant des marqueurs génétiques de type microsatellites (SSR).

### 1.1. L'hévéa et son amélioration génétique

Le genre *Hevea* est un complexe de 10 espèces (Schultes, 1990) largement interfertiles, cependant seule l'espèce *Hevea brasiliensis* ( $2n = 36$ ) produit suffisamment de latex pour être cultivée. L'hévéa est monoïque, préférentiellement allogame et très hétérozygote. La possibilité de multiplication végétative par greffage permet de stabiliser les meilleurs de ces génotypes et de les propager en tant que variétés clonales génétiquement homogènes. Les principaux objectifs de l'amélioration génétique sont le rendement en latex tout au long de la vie utile de l'arbre et la résistance à diverses maladies qui peuvent affecter les plantations en fonction du contexte écologique. Les programmes d'amélioration génétique sont majoritairement basés sur l'utilisation de la population Wickham dont l'origine génétique est restreinte mais qui a fait l'objet par le passé d'une sélection massale efficace sur les caractères d'intérêt agronomique. Des schémas de sélection à objectifs plus spécifiques (résistance à la maladie sud-américaine des feuilles ou production de bois) peuvent également être basés sur des populations sauvages Amazoniennes. La sélection d'individus élites s'effectue au sein de familles biparentales obtenues par pollinisation manuelle. Les clones d'une même génération sont comparés entre eux et à des témoins connus selon un dispositif en 2 étapes correspondant à 2 dispositifs expérimentaux différents : à petite échelle et sur des critères précoces puis à grande échelle et



sur des critères d'évolution de la production et du peuplement sur le moyen terme. Une durée d'environ 25 ans sépare la pollinisation manuelle du développement d'un nouveau clone (Nguyen Thi Hué 1997 ; Tran Thi Thuy Hoa 1998 ; RRIV 2010, Clément-Demange et al. 2007).

### 1.2. L'utilisation de marqueurs génétiques chez l'hévéa

Les marqueurs les plus couramment utilisés pour la réalisation de cartographies génétiques sont des marqueurs génétiques moléculaires situés au niveau de l'ADN. Les différents types de marqueurs moléculaires (RFLP, RAPD, AFLP, SSR et SNP) ont été développés chez l'hévéa et appliqués à des études génétiques, à l'exception des SNPs qui sont encore en phase de développement. A ce jour, les applications sont les suivantes : (1) Identification variétale et test de légitimité des croisements contrôlés (Leconte 1994) ; (2) Caractérisation de la diversité et gestion de la variabilité des ressources génétiques (Seguin 1999 ; Lekawipat et al. 2003 ; Le Guen et al. 2009) ; (3) Etude des flux de gènes en pollinisation naturelle (Blanc et al. 2001) ; (4) Analyse du génome par la cartographie génétique (Lepinasse et al. 2000a ; Le Guen et al. 2011, Tran 2012, Souza et al 2013) ; (5) Analyse du déterminisme génétique et cartographie de QTLs de la résistance à la maladie sud-américaine des feuilles (Lepinasse et al. 2000b ; Le Guen et al. 2003 ; Le Guen et al. 2007 ; Le Guen et al. 2011), et de la croissance et de la production (Rattanawong et al. 2012, Souza et al. 2013). Cette application est désormais la plus importante : elle vise le développement de la sélection assistée par marqueurs chez l'hévéa.

Le Cirad a opté depuis une quinzaine d'années pour une stratégie de développement de marqueurs de type SSR basée sur la création de banques enrichies (Seguin et al. 1997 ; Le Guen et al. 2010). Les microsatellites ou SSRs combinent une bonne qualité génétique (codominance, multi-allélisme, neutralité, abondance dans le génome ...), une bonne qualité technique (automatisation, protocoles simplifiés, équipement assez facilement accessible). Le génotypage aux locus SSRs fait appel à la technique très robuste de la PCR, ainsi qu'à une électrophorèse pour séparer les fragments selon leur taille et un marquage chimique (nitrate d'argent, radioactivité, fluorescence) pour les visualiser.

L'identification et le développement de microsatellites a permis d'obtenir pour toutes les plantes des nombres croissants de marqueurs, hors des gènes exprimés : SSRs génomiques, et dans les gènes exprimés : EST-SSRs ou SSRs géniques (Varshney et al. 2005).

### 1.3. L'évolution des techniques de génotypage par microsatellites

Le Cirad a successivement employé 3 techniques de génotypage des microsatellites, très semblables pour ce qui concerne la mise en œuvre de la PCR, mais qui diffèrent notablement en ce qui concerne l'électrophorèse et la révélation des amplicons. La dernière en date, à

laquelle nous avons eu recours pendant le stage, utilise un équipement permettant l'électrophorèse capillaire (ABI3500xL) avec 4 fluorochromes pour révéler les fragments. Cette méthode a permis de réduire la technicité nécessaire, la quantité de travail du fait de l'automatisation, les risques d'erreur, et d'augmenter considérablement le nombre d'échantillons et les possibilités de multiplexage lors de la migration post-PCR. Elle offre un débit d'analyse rendant possible la réalisation rapide d'une carte génétique.

#### 1.4. La cartographie génétique

La première carte génétique de l'hévéa a été publiée par Lespinasse et al. (2000a). Elle a été suivie par d'autres travaux (Le Guen *et al.*, 2011 ; Tran 2012 ; Souza *et al.*, 2013). La cartographie génétique chez l'hévéa est permise par le développement d'un nombre croissant de marqueurs microsatellites sur cette espèce (plus de 500 actuellement).

L'établissement d'une carte de liaison génétique nécessite au départ une population expérimentale constituée par une descendance particulière. Chez les plantes allogames comportant des individus hétérozygotes, on utilise la descendance F1 issue du croisement de deux parents hétérozygotes (c'est le cas de l'hévéa). La cartographie génétique consiste à utiliser l'information de génotypage obtenue sur un grand nombre de marqueurs au sein d'une famille dans le but de constituer des groupes de liaison pour lesquels sont calculés l'ordre et la distance entre les marqueurs. Celle-ci est calculée en fonction du taux de recombinaison observé entre ces marqueurs à l'issue de la méiose chez l'individu parental. Le taux de recombinaison, qui dépend de la fréquence des crossing-over, augmente avec la distance entre les deux locus. L'estimation du taux de recombinaison fournit donc une estimation de la distance entre deux locus. Un taux de recombinaison de 50% signifie qu'il y a indépendance complète. C'est le cas lorsque les locus sont situés sur des groupes de liaison différents ou lorsqu'ils sont éloignés sur un même groupe de liaison. Après avoir génotypé les descendants d'une population pour un ensemble de marqueurs, on teste d'abord les ratios de ségrégation de chaque marqueur afin de détecter les cas de distorsion de ségrégation, puis on étudie tous les couples de locus-marqueurs possibles afin de tester la distribution des gamètes parentaux et recombinés (de Vienne 1995). Les écarts aux effectifs théoriques peuvent être dus à une liaison génétique entre les deux locus ou à une distorsion sur un locus. Le test du  $\chi^2$  est adapté, mais les logiciels de cartographie utilisent plutôt la méthode du maximum de vraisemblance et la statistique du LOD score. Sur la base des liaisons observées, les logiciels de cartographie regroupent d'abord les marqueurs en groupes de liaison, puis ordonnent les marqueurs sur chaque groupe, estiment les taux de recombinaison entre marqueurs adjacents et convertissent ces taux en distances génétiques.

L'estimation des taux de recombinaison est obtenue par la méthode statistique du maximum de vraisemblance, réalisée automatiquement par les logiciels de cartographie génétique. Il faut convertir ces taux de recombinaison en distances génétiques pour pouvoir additionner les longueurs des segments et estimer la longueur totale de chaque groupe de liaison. Les cartes génétiques construites chez l'hévéa jusqu'ici ont utilisé la fonction de Kosambi qui, à taux de recombinaison égal, donne des distances plus réduites que la distance de Haldane. Une carte génétique est dite « saturée » quand chaque locus du génome est lié génétiquement à un marqueur. Cela suppose que l'on a trouvé autant de groupes de liaison que de chromosomes sur le génome haploïde de la plante, et que les marqueurs positionnés sur chaque groupe couvrent toute la longueur du groupe de liaison jusqu'à ses deux télomères.

### 1.5. La détection de QTLs

Une fois qu'on a établi une carte génétique saturée, on peut analyser si une ou plusieurs régions de cette carte sont statistiquement corrélées à l'expression d'un caractère d'intérêt. Il faut pour cela disposer d'une mesure fiable de ce caractère sur l'ensemble des individus du croisement étudié ayant servi à établir la carte.

Le principe de la détection de QTLs consiste, pour chaque locus marqueur, à répartir les individus du croisement étudié en classes génotypiques en fonction de la composition allélique des parents pour ce locus, et de comparer les valeurs moyennes de ces classes génotypiques pour le caractère d'intérêt. Une différence significative indique une liaison génétique entre le locus étudié et une région génomique participant à l'expression du caractère, appelée QTL. Plus le locus marqueur et le QTL sont proches physiquement, plus ils sont liés génétiquement et plus les différences de valeurs entre classes génotypiques sont importantes.

Pour tester la présence de QTLs nous avons utilisé une méthode paramétrique dite « Interval Mapping » (IM) pour laquelle le caractère étudié doit avoir une distribution normale. Cette méthode détermine pour chaque position d'un intervalle entre deux marqueurs successifs la probabilité de présence d'un QTL en fonction des informations aux marqueurs. La méthode utilisée est celle de l'estimation du maximum de vraisemblance et la statistique calculée est le LOD score. Le génome est scanné sur toute son étendue avec un pas de 1 cM. Une valeur seuil de significativité du LOD score est déterminée par une méthode de permutations. Il est donc possible d'établir pour chaque groupe de liaison une représentation graphique de la valeur du LOD. Un « pic » de LOD score supérieur au seuil de significativité traduit la présence proche d'un QTL participant à l'expression du caractère. Plus ce pic est élevé, plus la participation du QTL à l'expression du caractère est importante.

## 2. Résultats

### 2.1. Rappel de résultats antérieurs et synthèse des résultats nouvellement acquis

Un travail réalisé en 2012 avait initié la cartographie de la famille PBRIC (PB260 x RRIC100) sur un groupe de 190 descendants présents au Vietnam (Rriv) et de positionner 132 marqueurs SSR sur les 18 groupes de liaison de la carte de référence Cirad de l'hévéa (Tran, 2012).

Au cours de ce stage de M2, la cartographie a été poursuivie sur ce groupe de descendants, permettant le positionnement de 200 marqueurs supplémentaires, puis une approche nouvelle a permis de cartographier simultanément deux familles : PBRIM (PB260 x RRIM600) et PBRIC par le génotypage de deux groupes de 190 descendants présents dans des essais de phénotypage en champ en Côte d'Ivoire (Sogb) et au Nigéria (Orel) respectivement. Les deux familles ayant une mère commune, on a obtenu, par regroupement de l'ensemble des résultats de génotypage : une carte PBRIM portant sur 190 individus, une carte PBRIC portant sur 380 individus, et 3 cartes parentales avec 190 individus pour le clone paternel RRIM600, 380 individus pour le clone paternel RRIC100, et 570 individus pour le clone maternel PB260.

### 2.2. Extraction d'ADN

La quantité d'ADN obtenue par échantillon variait de 14000 à 268000 ng sur l'ensemble des 380 échantillons des deux familles. Ces quantités étaient suffisantes pour le génotypage d'environ 400 marqueurs ( $25 \text{ ng/marqueur} \times 400 \text{ marqueurs} = 10000 \text{ ng}$ ).

### 2.3. Test de légitimité des descendants

Avec 8 marqueurs SSR, on a vérifié la légitimité de 240 descendants par famille, choisis parmi ceux qui avaient été mis en essai en champ à la Sogb et à Orel. On a détecté 18 individus illégitimes dans la famille PBRIM (7,5 %), et 9 illégitimes pour PBRIC (3,7 %).

### 2.4. Génotypage des individus des deux descendance

Parmi les descendants légitimes de PBRIM (222) et de PBRIC (231), on a sélectionné 190 individus par famille, en choisissant ceux qui avaient le plus grand nombre de répétitions dans les essais en champ afin d'obtenir la meilleure précision possible dans la détection de QTLs.

Le tri parmi 494 marqueurs SSR disponibles au Cirad a permis d'obtenir 403 marqueurs polymorphes sur au moins un des trois parents des deux familles (Tran, 2012). Le génotypage de ces 403 marqueurs sur les 190 descendants choisis de chaque famille donne la répartition des effectifs de marqueurs pour chacune des deux familles et pour les 5 types de ségrégation possibles, et les effectifs pour les 3 parents (tableau 1).

**Tableau 1.** Répartition des effectifs de s 403 marqueurs SSRs par type de ségrégation, par famille et par parent.

Ordre	Type de ségrégation	Génotypages possibles	PBRIM	PBRIC
1	<abxcd>	ac, ad, bc, bd	38	67
2	<efxeg>	ee, ef, eg, fg	82	89
3	<hkxhk>	hh, hk, kk, h-, k-	36	21
4	<lmxll>	ll, lm	95	71
5	<nnxnp>	nn, np	57	93
	<b>Total</b>		<b>308</b>	<b>341</b>
	<i>Polymorphes chez PB260</i>		251	248
	<i>Polymorphes chez RRIM600</i>		213	
	<i>Polymorphes chez RRIC100</i>			270

## 2.5. Cartographie génétique

Avec le logiciel Joinmap, on a traité une matrice de données de génotypage de 190 individus x 308 marqueurs (58520 données bi-alléliques) avec 1369 données manquantes (2,3 %) pour PBRIM, et une matrice de 190 individus x 341 marqueurs (64790 données) avec 1683 données manquantes (2,6 %) pour PBRIC. Les marqueurs portent dans le préfixe de leur nom une hypothèse concernant leur appartenance à un groupe de liaison : ces marqueurs ont déjà été cartographiés sur d'autres familles, ou ont fait l'objet de tests de « bin-mapping » sur un petit nombre d'individus dans une famille donnée. Ces hypothèses *a priori* sur l'appartenance des marqueurs aux groupes de liaison de référence ne sont pas utilisées dans l'analyse mais servent de moyen de contrôle *a posteriori*.

### 2.5.1. Cartographie de la famille PBRIM

Une analyse « simple locus » a permis de rechercher les distorsions de ségrégation. Le marqueur g16M421 présentait une forte distorsion car il avait une ségrégation <hk x hk> interprétée par défaut par le logiciel selon un modèle de classification co-dominant (hh, hk, kk) avec le ratio 1 : 2 : 1. On a corrigé l'erreur en indiquant au logiciel qu'il s'agissait d'une classification dominante (h-, kk) avec le ratio 3 : 1. Ensuite on a procédé au regroupement des marqueurs en groupes de liaison. Au seuil de LOD = 5, on a obtenu les 18 groupes de liaison attendus, correspondant aux 18 chromosomes du génome haploïde de l'hévéa, avec une affectation conforme de tous les marqueurs à leurs groupes de liaison supposés. On a enfin ordonné les marqueurs sur chaque groupe de liaison. Quatre marqueurs ont été exclus de la

carte, soit parce que leur présence perturbait l'ordre des marqueurs voisins, soit parce que leur positionnement était douteux. Les caractéristiques de la carte PBRIM et de chaque groupe de liaison sont présentées dans le tableau 2 et la figure 1. La longueur totale de la carte est de 2004 cM, avec 304 marqueurs. Les longueurs des groupes de liaison varient entre 71 cM (groupe g12) et 147 cM (groupe g8). La longueur moyenne des segments compris entre deux marqueurs consécutifs est de 7.0 cM, mais on trouve 18 segments de longueur supérieure à 20 cM (Figure 1). Les longueurs des 5 segments les plus grands varient entre 32 et 44 cM.

#### 2.5.2. Cartographie de la famille PBRIC

On a d'abord cartographié cette famille avec les données de génotypage d'un groupe de descendants présent au Nigéria. L'analyse de ségrégation des 341 marqueurs a permis de détecter 3 marqueurs fortement distordus ; pour l'un d'entre eux, on a modifié l'option par défaut sur la classification (hh, hk, kk) et indiqué la classification adéquate (h-, kk). Les deux autres marqueurs distordus ont été exclus. Avec une matrice de données de génotypage de 339 marqueurs x 190 individus, en utilisant un seuil de LOD = 4, on a obtenu les 18 groupes de liaison attendus avec une confirmation de l'appartenance de tous les marqueurs aux groupes de liaison supposés. Après la mise en ordre des marqueurs sur chaque groupe de liaison, on a obtenu une carte de longueur totale égale à 1758 cM avec 338 marqueurs (le marqueur g5hmg1 jugé trop mal placé sur le groupe 5 a été éliminé). L'intervalle moyen des marqueurs adjacents est de 5.5 cM. Les caractéristiques de cette carte PBRIC réalisée avec le groupe de descendants du Nigéria sont données dans le tableau 2 et la figure 2.

Pour améliorer la précision de cette carte, les données ont ensuite été regroupées avec celles obtenues dans l'étude précédente (Tran 2012) sur des individus du même croisement présents au Vietnam. La carte des deux échantillons regroupés de la famille PBRIC, avec 354 marqueurs et 380 individus, présente une longueur de 1916 cM, un peu supérieure à celles des cartes des descendants du Nigeria (1758 cM) et du Vietnam (1869 cM) réalisées chacune avec 190 individus. L'intervalle moyen entre deux marqueurs consécutifs est de 5,7 cM. On trouve 12 segments ayant une longueur supérieure à 20 cM, et les longueurs des 5 segments les plus grands varient entre 23 et 28 cM.

#### 2.5.3. Cartographie des 3 parents

(Les caractéristiques des 3 cartes parentales sont présentées dans le tableau 3 et annexe 1)

Carte de RRIM600 : On a traité une matrice de 190 individus x 213 marqueurs avec 1034 données manquantes. Joinmap a permis le regroupement des 18 groupes de liaison

attendus avec les marqueurs correctement localisés sur leurs groupes supposés. La longueur totale obtenue était de 1829 cM avec 213 marqueurs. La longueur de chaque groupe variait de 53 à 147 cM, avec une longueur de 9.4 cM pour l'intervalle moyen entre deux marqueurs consécutifs. Le nombre de marqueurs par groupe de liaison variait de 7 (g17) à 20 (g10).

Carte de RRIC100: On avait deux échantillons provenant du Vietnam et du Nigéria pour la famille PBRIC. Pour chaque échantillon, on a construit une carte génétique de RRIC100 comme indiqué ci-dessus pour la cartographie de RRIM600. A partir de deux fichiers «LOC», on a combiné sur Joinmap deux fichiers «LOC» pour réaliser une carte synthétique de RRIC100. On a obtenu les 18 groupes de liaison attendus. La longueur totale était de 1822 cM avec 283 marqueurs, et la distance moyenne entre deux marqueurs adjacents était de 6,9 cM. Le groupe de liaison le plus long était le groupe 10 (137 cM), et le plus court était le groupe 17 (69 cM). Le nombre de marqueurs par groupe variait de 7 (groupe 12) à 37 (groupe 10). L'ordre des marqueurs sur chaque groupe de liaison était bien conservé par comparaison avec celui de la carte de la famille PBRIC

Carte de PB260: Ce parent possédait le plus grand nombre d'individus (570 individus) issus des trois échantillons provenant des deux familles, avec 190 individus pour la famille PBRIM de Côte d'Ivoire, et 380 individus pour la famille PBRIC (190 du Vietnam et 190 du Nigéria). Pour chaque échantillon, on a sélectionné les marqueurs polymorphes sur PB260, puis on a regroupé les trois échantillons pour produire la carte. On a obtenu les 18 groupes de liaison attendus, avec une confirmation de l'appartenance des marqueurs à leurs groupes de liaison supposés. La longueur totale obtenue pour cette carte est de 1899 cM, avec 271 marqueurs. La longueur de chaque groupe variait de 66 cM (groupe 2) à 153 cM (groupe 10), avec un intervalle moyen entre marqueurs flanquants de 7,5 cM. Le nombre de marqueurs sur chaque groupe de liaison variait de 9 (groupe 12) à 31 (groupe 10). L'ordre des marqueurs sur chacun des 18 groupes de liaison de la carte de PB260 réalisée avec 570 individus était très voisin de ceux qu'on avait observés sur les cartes correspondant aux trois échantillons simples et aux deux familles.

La figure 3, constituée à partir de données du tableau 3, illustre les variations des distances génétiques des groupes de liaison des 3 parents.

La figure 4 présente les alignements des marqueurs entre les 5 cartes génétiques des 2 familles et des 3 parents pour le groupe de liaison g10 qui présente à la fois la plus grande longueur moyenne, le plus grand nombre de marqueurs et la plus faible distance moyenne des segments entre marqueurs consécutifs.

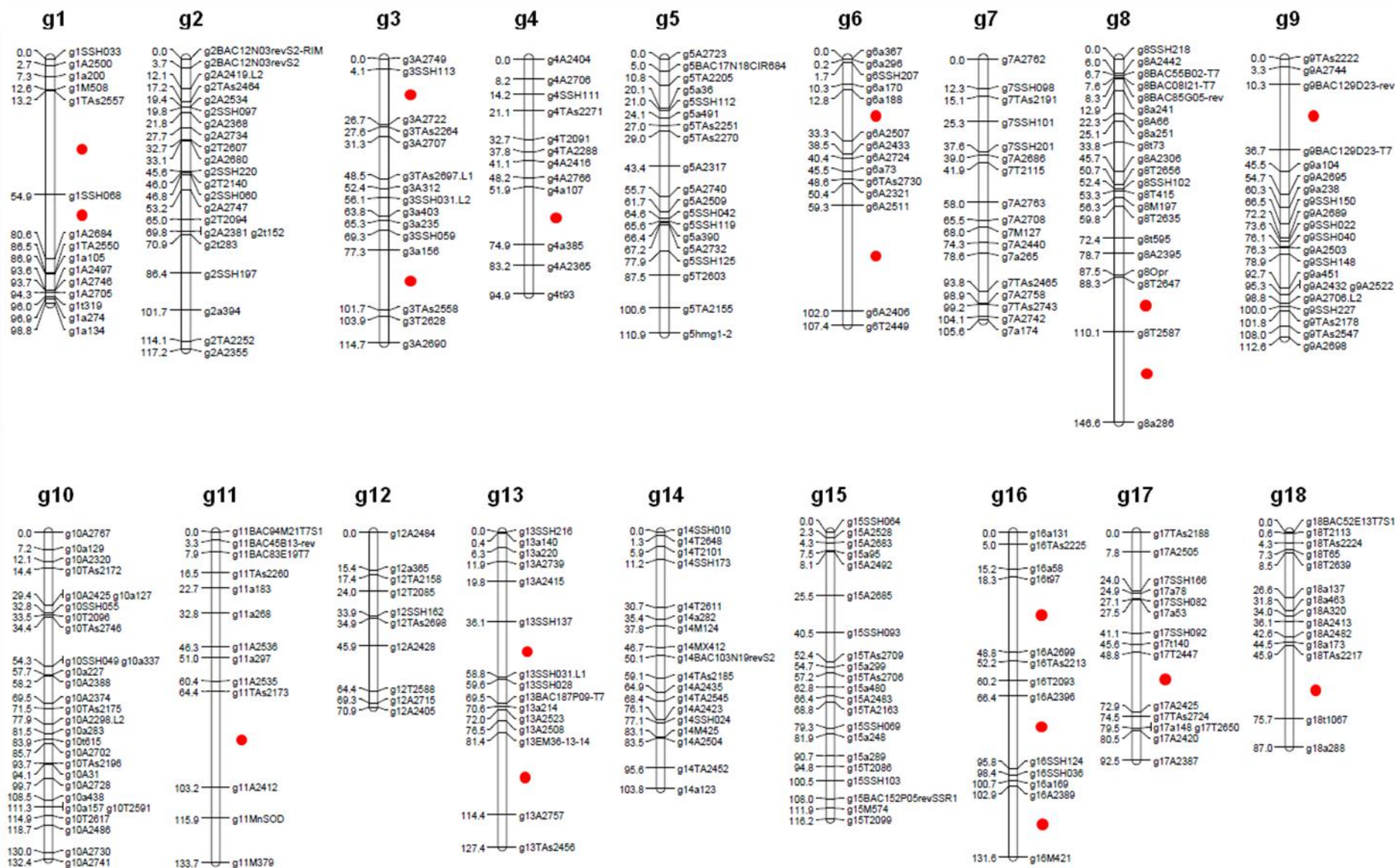
**Tableau 2.** Répartition des 403 marqueurs et longueurs en cM des groupes de liaison sur les différentes cartes des deux familles.

Groupe de liaison	Carte de la famille PBRIM (190 ind)			Carte de la famille PBRIC									Marqueurs communs à deux familles
				PBRIC Nigéria (190 ind)			PBRIC Vietnam (190 ind)			PBRIC regroupé (380 ind)			
	Nbre loci	Longueur (cM)	Intervalle (cM)	Nbre loci	Longueur (cM)	Intervalle (cM)	Nbre loci	Longueur (cM)	Intervalle (cM)	Nbre loci	Longueur (cM)	Intervalle (cM)	
Groupe 1	15	99	7.1	17	85	5.3	16	86	5.7	18	89	5.2	15
Groupe 2	22	117	5.6	21	109	5.4	23	111	5.0	25	111	4.6	20
Groupe 3	15	115	8.2	21	100	5.0	23	102	4.6	23	101	4.6	15
Groupe 4	12	95	8.6	15	101	7.2	13	92	7.7	15	97	6.9	12
Groupe 5	19	111	6.2	23	75	3.4	23	125	5.7	23	120	5.5	18
Groupe 6	14	107	8.3	16	71	4.8	17	116	7.3	16	113	7.5	14
Groupe 7	17	106	6.6	20	104	5.5	20	103	5.4	21	104	5.2	17
Groupe 8	21	147	7.3	22	114	5.4	21	111	5.6	22	115	5.5	17
Groupe 9	21	113	5.6	22	102	4.9	22	97	4.6	23	101	4.6	20
Groupe 10	29	132	4.7	38	134	3.6	36	136	3.9	42	138	3.4	29
Groupe 11	13	134	11.1	15	126	9.0	13	95	7.9	15	117	8.4	14
Groupe 12	10	71	7.9	9	51	6.3	10	75	8.3	10	76	8.5	9
Groupe 13	15	127	9.1	15	122	8.7	14	125	9.6	15	123	8.8	15
Groupe 14	18	104	6.1	19	87	4.8	18	110	6.5	19	107	5.9	18
Groupe 15	21	116	5.8	22	146	7.0	19	106	5.9	21	126	6.3	18
Groupe 16	13	132	11.0	13	110	9.1	12	110	10.0	13	111	9.2	12
Groupe 17	15	93	6.6	14	68	5.2	15	69	4.9	15	69	4.9	13
Groupe 18	14	87	6.7	16	55	3.7	17	100	6.3	18	101	6.0	13
Total	304	2004		338	1758		332	1869		354	1916		289
Intervalle moyen			7.0			5.5			6.0			5.7	



**Tableau 3.** Répartition des marqueurs et longueur de groupe de liaison sur trois parents.

Groupe de liaison	RRIM600 (190 ind)			RRIC100 (380 ind)			PB260 (570 ind)			Marqueurs communs à 3 parents
	Nbre loci	Longueur (cM)	Intervalle (cM)	Nbre loci	Longueur (cM)	Intervalle (cM)	Nbre loci	Longueur (cM)	Intervalle (cM)	
Groupe 1	12	100	9.1	10	82	9.2	16	90	6.0	10
Groupe 2	17	115	7.2	20	110	5.8	19	66	3.7	17
Groupe 3	8	112	16.0	17	99	6.2	14	117	9.0	8
Groupe 4	9	91	11.4	13	95	7.9	12	96	8.7	9
Groupe 5	16	109	7.3	22	109	5.2	13	86	7.2	13
Groupe 6	8	59	8.4	10	100	11.2	14	117	9.0	8
Groupe 7	12	104	9.5	21	105	5.2	16	109	7.2	12
Groupe 8	18	147	8.6	21	118	5.9	11	101	10.1	11
Groupe 9	9	86	10.7	16	89	6.0	19	115	6.4	9
Groupe 10	20	119	6.3	37	137	3.8	31	153	5.1	20
Groupe 11	9	129	16.1	12	117	10.6	12	114	10.4	9
Groupe 12	8	53	7.6	7	74	12.3	9	76	9.5	7
Groupe 13	14	127	9.8	10	118	13.1	15	126	9.0	10
Groupe 14	13	93	7.7	11	86	8.6	17	115	7.2	11
Groupe 15	11	103	10.3	16	111	7.4	17	123	7.7	11
Groupe 16	10	130	14.4	9	102	12.7	12	108	9.8	9
Groupe 17	7	67	11.1	14	69	5.3	14	69	5.3	7
Groupe 18	12	86	7.8	17	102	6.4	10	117	13.0	10
<b>Total</b>	<b>213</b>	<b>1829</b>		<b>283</b>	<b>1822</b>		<b>271</b>	<b>1899</b>		<b>191</b>
<i>Intervalle moyen</i>			9.4			6.9			7.5	



**Figure 1.** Carte génétique de la famille PB260 x RRIM600 (190 individus, 304 marqueurs, (●) segment sans marqueur supérieur 20 cM)

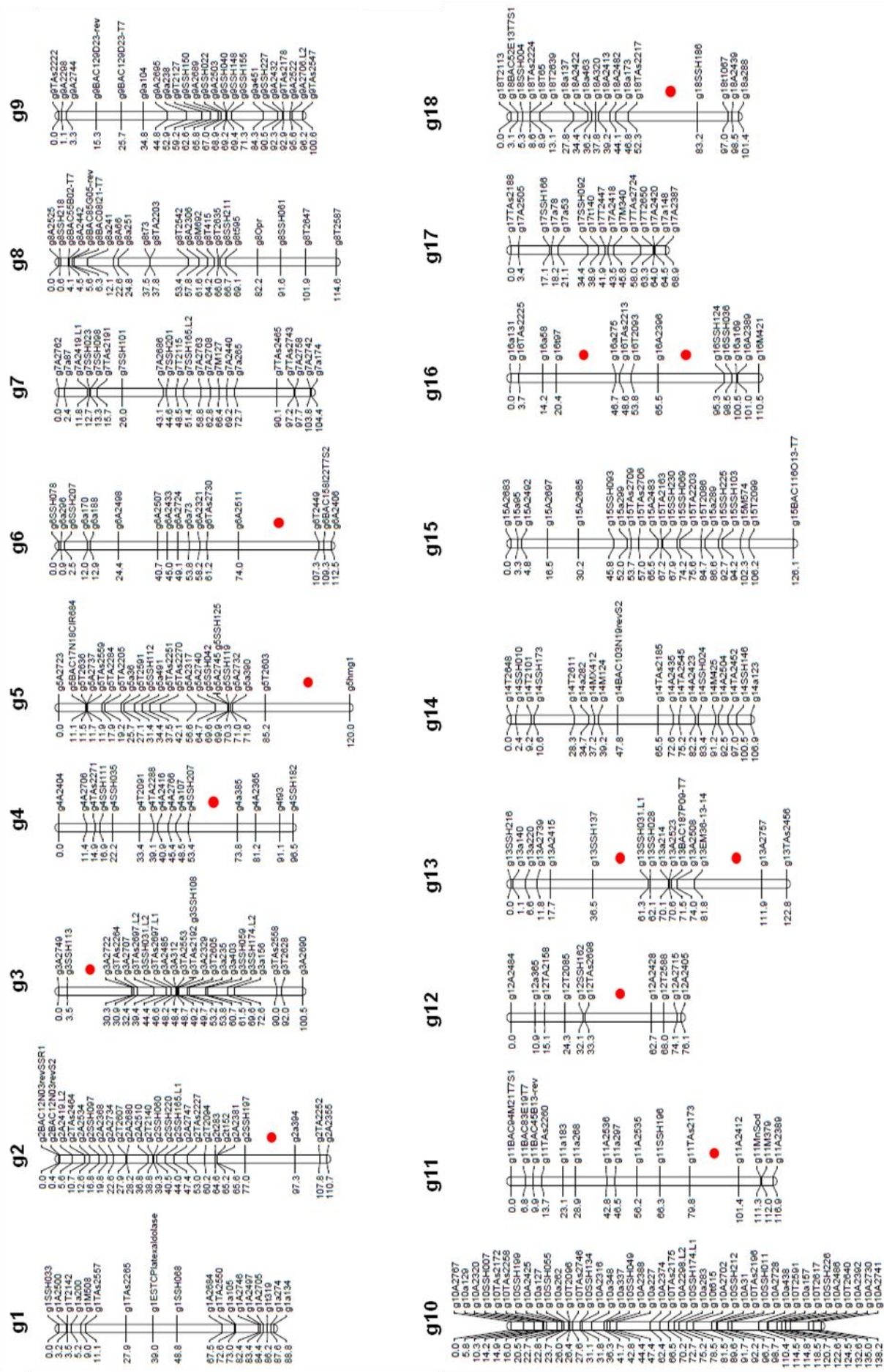
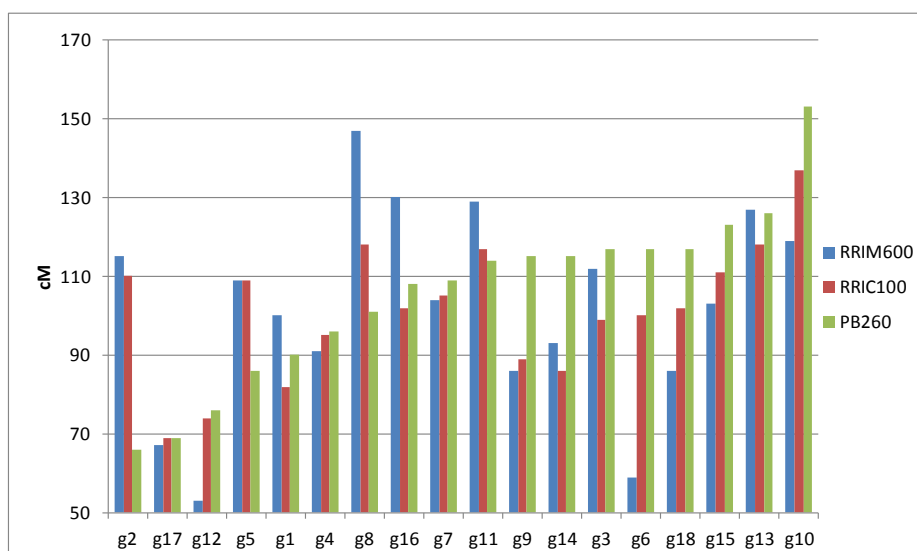
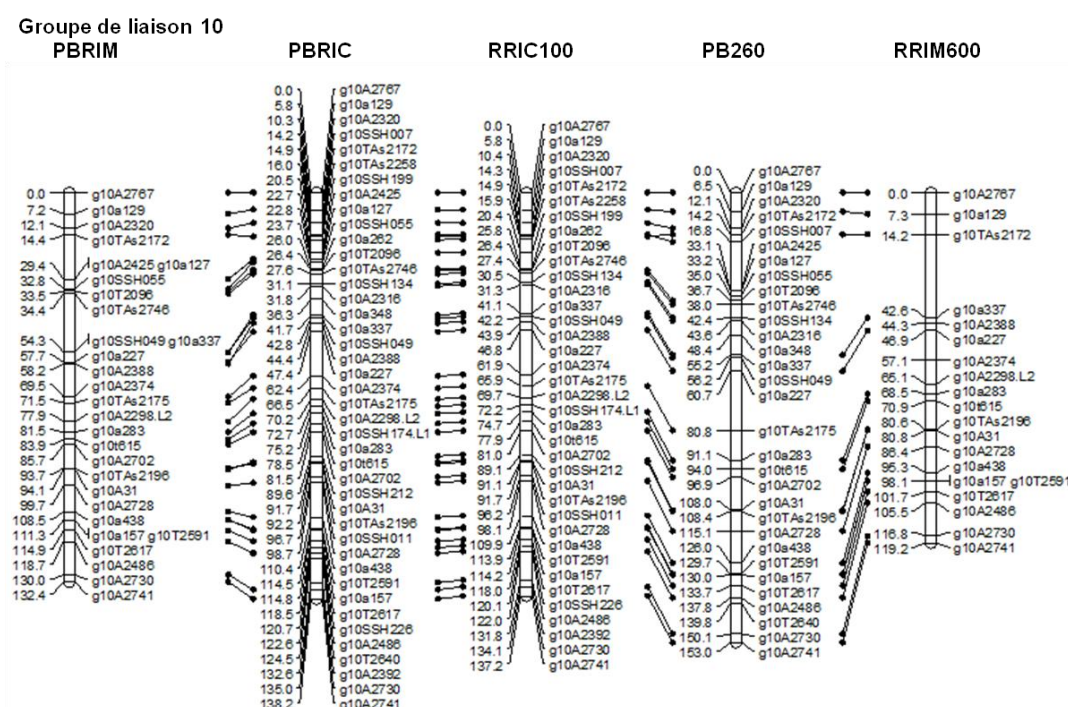


Figure 2. Carte génétique de la famille PB260 x RRIC100 (380 individus, 354 marqueurs, (●) segment sans marqueur supérieur 20 cM





**Figure 3.** Variations des distances génétiques des groupes de liaison des 3 parents (cM).



**Figure 4.** Alignement des marqueurs du groupe g10 entre les cartes des 2 familles et des 3 parents.

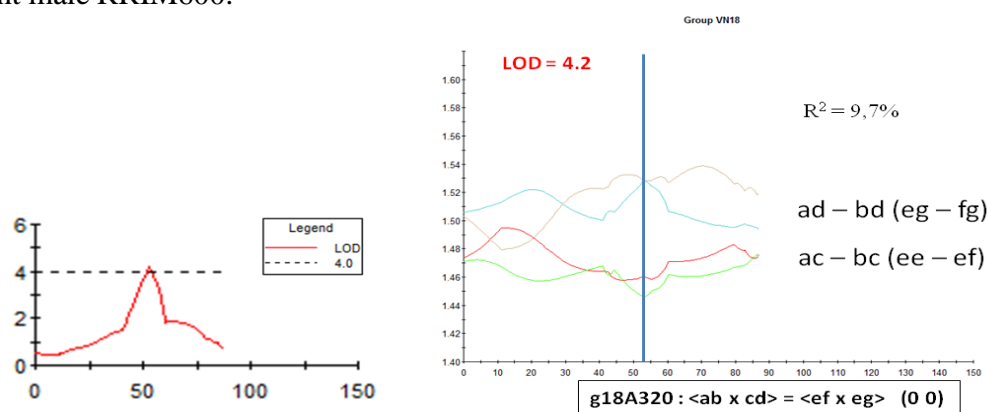
## 2.6. Détection de QTLs

Avec les données de génotypage des descendants expérimentés en champ et les cartes des deux familles, on a réalisé une tentative de détection de QTLs sur les premières mesures de croissance des arbres plantés en 2012. Les valeurs génétiques des descendants pour ces caractères de croissance ont été estimées avec le logiciel SAS selon un modèle mixte (estimations de type BLUP). Les héritabilités au sens large des moyennes phénotypiques pour

les caractères D4 et D9 mesurées sur PBRIM (Côte d’Ivoire) et pour le caractère D7 mesuré sur PBRIC (Nigéria) variaient selon le nombre de répétitions ; pour 10 répétitions, on a obtenu les estimations d’héritabilité suivantes :  $h^2 = 0.26$  pour D4 et  $h^2 = 0.20$  pour D9 sur PBRIM, et  $h^2 = 0.22$  pour D7 sur PBRIC. Entre les estimations Blup des génotypes, pour la famille PBRIM, le caractère D4 variait de 0.753 à 0.849 cm, et le caractère D9 variait de 1.408 à 1.557 cm ; pour la famille PBRIC, le caractère D7 variait de 1.215 à 1.384 cm.

Les seuils de signification du LOD score établis par des tests de permutation sur 2000 itérations variaient entre 4.2 et 4.3 (un risque global  $\alpha$  sur l’ensemble du génome  $\alpha = 0.05$ ).

Pour le caractère de diamètre des tiges D4 mesuré à 4 mois à la Sogb (famille PBRIM), le LOD score maximum était de 3.3, mais ce résultat était dû à l’effet d’un seul marqueur. Sur le groupe g18, un LOD = 3.2 a été obtenu à la position du marqueur g18A463 ; dans ce cas, bien que le LOD soit inférieur au seuil de signification, on observe également l’effet sur les marqueurs voisins. Pour le caractère D9 mesuré à 9 mois, un LOD = 4.2, à la limite du seuil de signification, a été obtenu à la position du marqueur g18A320 (position 53.0 cM) proche de celle du marqueur g18A463 (position 55.2 cM). Ce QTL expliquait 9.7 % de l’ensemble de la variance phénotypique du caractère D9. Pour ce caractère D9 et pour le marqueur g18A320 qui est le plus proche du QTL nommé g18-53 (locus situé à la position 53 cM du groupe 18), une analyse de variance testant les effets de chaque parent et de l’interaction éventuelle a été réalisée sur les 2 x 2 classes constituées par les allèles de PB260 et de RRIM600 (Fig 5). Elle montre que l’effet du QTL est dû exclusivement à un effet hautement significatif entre les deux allèles du parent mâle RRIM600.



**Figure 5.** Détection de QTL de D9 (diamètre à 9 mois) sur le groupe 18 de la famille PBRIM

Pour le caractère D7 mesuré à 7 mois à Orel (famille PBRIC), le LOD score maximum était de 2.95, nettement inférieur au seuil de signification ; il n’y a pas donc un QTL (annexe 4).

### 3. Discussion

Le travail de ce stage de Master 2 a porté sur le génotypage de 3 groupes de 190 individus représentant 3 parents et 2 familles, et la cartographie de plus de 300 marqueurs SSR par famille. La localisation sur les différents groupes de liaison des 403 marqueurs utilisés est désormais bien établie. Les résultats de génotypage et de cartographie associés aux premières mesures de croissance des arbres en champ ont permis la détection d'un premier QTL.

#### 3.1. Longueur des cartes

Les longueurs totales des cartes obtenues ici varient de 1758 cM (PBRIC Nigéria, 190 individus) à 2004 cM (PBRIM, 190 individus). Ces longueurs sont un peu inférieures à celle qui avait été obtenues par Lespinasse et al (2000) sur la famille PB260 x RO38 (longueur de 2144 cM pour 106 individus et 717 marqueurs) ou à celle qui avait été obtenue par Rattanawong et al (2008) sur la famille RRIM600 x PB217 (longueur de 2075 cM pour 196 individus et 427 marqueurs). En revanche, Le Guen et al (2011) obtenait une longueur totale plus importante pour la famille PB260 x MDF180 (2441 cM pour 351 individus et 383 marqueurs).

Concernant le parent PB260, on obtient ici une longueur de 1899 cM pour 570 individus et 271 marqueurs, ce qui est assez comparable au résultat de Lespinasse et al (2000) qui était de 2000 cM pour 106 individus et 259 marqueurs ; en revanche, Le Guen et al (2011) obtenait une longueur de 2648 cM pour 351 individus et 249 marqueurs. Concernant le parent RRIM600, on obtient ici une longueur de 2004 cM avec 304 marqueurs, un peu supérieure à la longueur de 1848 cM obtenue avec 326 marqueurs par Rattanawong et al (2008). Les deux cartes de PBRIC réalisées ici avec deux groupes distincts de 190 descendants échantillons, un même jeu de marqueurs et une même méthode ont des longueurs un peu différentes, 1758 cM pour les descendants du Nigéria et 1869 cM pour les descendants du Vietnam, et cette différence ne peut être due qu'à un effet d'échantillonnage des descendants.

#### 3.2. Alignements des marqueurs sur les différentes cartes

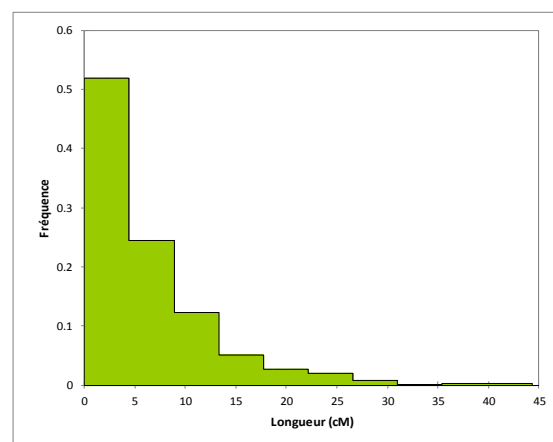
Pour les 18 groupes de liaison, la comparaison entre elles des 3 cartes parentales montre une forte stabilité de l'ordre des marqueurs communs à au moins deux cartes (annexe 1).

#### 3.3. Variabilité des longueurs des segments entre deux marqueurs consécutifs

L'examen des 3 cartes parentales montre, sur différents groupes de liaison, des zones pauvres en marqueurs, conduisant parfois à l'absence d'une extrémité d'un groupe pour un parent par comparaison avec les deux autres parents.

La figure 6 met en évidence l'irrégularité de la répartition des marqueurs sur le génome. Tandis que plus de 50 % des segments ont une longueur inférieure à 5 cM, 5 % du nombre de segments ont une longueur comprise entre 20 et 44 cM.

**Figure 6.** Variabilité des longueurs des segments compris entre deux marqueurs consécutifs. Données regroupées des cartes des deux familles PBRIM et PBRIC.



Concernant ces grands segments, on peut selon une première hypothèse, considérer qu'ils correspondent à des régions à haute fréquence de recombinaison où les distances génétiques surestiment les distances physiques. Selon une seconde hypothèse, on peut supposer que ces régions seraient pauvres en marqueurs microsatellites. Selon une troisième hypothèse, des marqueurs polymorphes existeraient dans ces régions mais ne feraient pas partie de la collection disponible actuellement au Cirad.

Pour réaliser la carte génétique de la famille RRIM600 x PB217 (Rattanawong et al, 2008), on avait ajouté 192 marqueurs AFLP à 235 marqueurs SSR, permettant de n'avoir plus que 6 segments de plus de 20 cM. D'autres types de marqueurs pourraient être envisagés tels que les DArT ou les SNPs pour améliorer la couverture du génome par les marqueurs. Cependant, l'utilisation d'un seul type de marquage moléculaire SSR a considérablement facilité la rapidité d'exécution du génotypage des familles PBRIM et PBRIC. Si l'on veut améliorer la couverture du génome par l'emploi des seuls marqueurs microsatellites afin de maintenir une bonne productivité du travail de génotypage, on devrait donc chercher à enrichir la collection actuelle de ces marqueurs.

### 3.4. Détection du QTL g18-53 pour le caractère D9 sur la famille PBRIM

Le QTL g18-53 détecté sur la carte PBRIM, porte sur une mesure de diamètre des tiges de jeunes plants en champ, réalisée à un stade très précoce (9 mois). Il n'est qu'à la limite de signification et il explique moins de 10 % de la variance observée, ce qui reste modeste pour une application à la sélection. On peut cependant espérer obtenir des effets plus marqués pour les mesures ultérieures, au vu de résultats obtenus sur la famille RRIM600 x PB217 en

Thaïlande où l'effet d'un QTL de croissance, détecté à partir de 18 mois avec d'abord un effet faible, s'était accru avec le temps pour atteindre à 5 ans un  $R^2$  égal à 30 % de la variance phénotypique totale du caractère (Rattanawong et al. 2008).

La technique actuelle de l'électrophorèse capillaire pour le génotypage de marqueurs SSR offre un débit qui devrait permettre une mise en œuvre effective de la sélection assistée par marqueurs mais dans un cadre limité, celui de la sélection intra-familiale, nécessitant la cartographie de chaque famille étudiée sur la base du génotypage de 190 individus. On ne dispose pas d'information précise pour justifier le nombre retenu de 190 descendants par famille pour la détection de QTLs. Dans le cadre de ce stage, nous avons fait des simulations sur différents échantillons qui ont montré qu'on aurait pu obtenir des cartes satisfaisantes avec des effectifs plus faibles. Cependant le nombre d'individus influence beaucoup la précision de la détection des QTLs. Des recherches de QTLs avec 100 à 120 individus par famille F1, portant par exemple sur le taro (Quero-Garcia et al, 2006) ou le pommier (Segura et al, 2006), ont conduit à certains résultats intéressants. Il serait donc utile de ré-analyser les données disponibles sur des effectifs plus faibles pour apprécier la perte de précision qui en résulterait. L'expérience acquise avec la famille RRIM600 x PB217 (Rattanawong et al, 2008) indique que cet effectif permet une détection efficace des QTLs expliquant au moins 5 % de la variance phénotypique des caractères étudiés chez l'hévéa.

La détection de QTLs sur des familles F1 d'hévéa permettra une accumulation de connaissances sur le déterminisme génétique des caractères exprimés dans différents contextes écologiques. Elle permettra aussi d'améliorer la précision des estimations de valeurs génétiques en ajoutant des informations génotypiques aux données phénotypiques mesurées en champ, et de contre-sélectionner de façon ciblée les allèles défavorables aux QTLs, ce que ne permet pas une sélection phénotypique portant sur la globalité d'un caractère. De plus, une fois la détection de QTLs faite sur un échantillon familial relativement restreint (190 individus), il deviendra possible de réaliser une pré-sélection, sur la base des marqueurs seuls, sur un échantillon élargi de la famille (par exemple sur 1000 descendants) et de n'installer en champ pour la poursuite de la sélection que les génotypes déjà retenus par cette pré-sélection génotypique. On réalisera ainsi une sélection assistée par marqueurs (SAM) au niveau du pre-breeding intra-familial.

Les perspectives pour aller plus loin dans le développement de la SAM reposent sur le développement actuel de marqueurs SNP chez l'hévéa, en vue d'étudier les déséquilibres de liaison au sein de la population d'amélioration Wickham et d'explorer les possibilités de la génétique d'association pour une sélection à l'échelle de la population toute entière.



## **4. Matériel et Méthodes**

### **4.1. Matériel végétal**

Deux familles F1 ont été réalisées au Vietnam (Rriv) par pollinisation manuelle en 2011 : PB260 x RRIM600 (PBRIM), et PB260 x RRIC100 (PBRIC). Une partie des graines a été conservée pour un phénotypage en champ au Vietnam (Rriv). L'autre partie a été expédiée en Côte d'Ivoire (famille VN1, plantation Sogb) et au Nigéria (famille VN2, plantation Orel) pour un phénotypage en champ. Après germination des graines et développement de plantules sur les 3 sites, trois groupes de folioles de ces génotypes ont été expédiées au Cirad pour l'extraction d'ADN et le génotypage.

### **4.2. Génotypage et cartographie**

Le tri des amorces de 494 marqueurs SSR disponibles au Cirad a révélé 403 marqueurs polymorphes sur au moins un des trois parents constitutifs des deux familles (Tran, 2012). Après avoir génotypé les 190 descendants PBRIC du Vietnam, on a génotypé en même temps les deux séries de 190 individus de Côte d'Ivoire et du Nigéria pour les deux familles et leurs trois parents sur une plaque 384 (Figure 7). De ce fait, certains marqueurs étaient parfois monomorphes pour l'une ou l'autre des deux familles.

L'électrophorèse capillaire réalisée avec le séquenceur ABI3500xL a permis le multiplexage de 12 marqueurs par migration (3 marqueurs x 4 couleurs de fluorochromes). Les réactions de PCR étaient effectuées en simplex et il fallait donc effectuer 12 PCR pour une migration. Le logiciel Genemapper utilisé pour la lecture des résultats, permet de déterminer la taille des allèles des 12 marqueurs pour chaque individu analysé.

Dans une population F1 entre parents hétérozygotes, on trouve au maximum 4 allèles pour chaque marqueur (2 allèles par parent) ce qui permet par recombinaison un maximum de 4 génotypes bi-alléliques parmi les descendants. Selon les marqueurs, il existe 5 types possibles de ségrégations, présentées dans le tableau 1. Après codage des allèles dans un fichier Excel, la matrice de données « individus x marqueurs » est analysée avec le logiciel JoinMap.

### **4.3. Dispositif expérimental de phénotypage en champ**

Deux essais de phénotypage ont été mis en place à la Sogb (famille PBRIM, 265 génotypes) et à Orel (famille PBRIC, 250 génotypes). Les plants greffés en sacs ont été plantés en champ en juillet 2012 à la Sogb, et en septembre 2012 à Orel. Ces essais ont été plantés à la densité de 1600 arbres par hectare (2.5 x 2.5 m) et installés pour une durée de 4 ans devant permettre des mesures de croissance, de production de latex, de teneur en saccharose du latex

(indicateur de tolérance à l'intensification de la saignée), et de tolérance à la maladie des feuilles *Corynespora cassiicola*. Le dispositif expérimental est constitué de blocs randomisés, incomplets et partiellement équilibrés, avec la mise en place de 11 blocs à la Sogb et 12 blocs à Orel, prenant en compte l'hétérogénéité apparente des terrains. Dans l'essai Sogb, le nombre d'arbres greffés par génotype varie de 3 à 18, avec au total 2961 arbres greffés pour 265 génotypes. Dans l'essai Orel, le nombre d'arbres greffés par génotype varie de 1 à 21, avec au total 2992 arbres greffés pour 250 génotypes.

#### 4.4. Détection de QTLs

Les premières mesures de croissance des arbres greffés (diamètre des greffons à 10 cm au dessus de la jonction porte-greffe/greffon) ont été réalisées à 4 mois (D4) et à 9 mois à la Sogb (D9), et à 7 mois (D7) à Orel. Pour ces 3 caractères, les estimations de type Blup des valeurs génétiques des génotypes (Henderson, 1984) ont été réalisées grâce au logiciel statistique SAS selon le modèle mixte suivant :

$$Y_{ijk} = \mu + b_i + G_j + \varepsilon_{ijk} \quad \left\{ \begin{array}{l} - \quad b : \text{blocs, effet fixe} \\ - \quad G : \text{génotype, effet aléatoire} \\ - \quad \varepsilon : \text{erreur résiduelle, effet aléatoire} \end{array} \right.$$

L'estimation Blup de la valeur génétique d'un génotype  $G_i$  est de la forme :

$$\widehat{G_i} = h^2 l_i * (P_i - \mu) \quad , \quad \text{avec} \quad h^2 l_i = \sigma^2 G / (\sigma^2 G + \sigma^2 \varepsilon / n_i)$$

( $P_i$  = moyenne des valeurs phénotypiques de  $G_i$  ;  $n_i$  = nombre de répétitions du génotype  $G_i$ ).

L'héritabilité  $h^2 l_i$  dépend du nombre de répétitions du génotype.

Ces estimations de valeurs génétiques, combinées avec les données de génotypage disponibles pour 190 individus de l'essai Sogb et 190 individus de l'essai Orel, et avec les cartes génétiques des deux familles, ont permis les premières tentatives de détection de QTLs sur ces essais. La détection de QTLs a été réalisée grâce au logiciel MapQTL5 (Van Ooijen, 2004) selon la méthode « Interval Mapping » adaptée à l'étude de caractère phénotypiques à distribution normale. Les seuils de signification du LOD score ont été estimés par la méthode de simulation numérique du test de permutation ; pour les 3 caractères.

**Figure 7.** Disposition de deux familles avec 190 individus + deux parents pour chaque famille sur une plaque 384 :

Chaque réaction PCR a besoin de 5 µl d'ADN à 5 ng/µl.

Il faut donc au minimum 5 µl x 400 marqueurs par échantillon = 2000 µl d'ADN à 5 ng/µl (ou 10000 ng)

Préparation de 4 racks (96 puits) d'ADN : Rack1 + Rack2 (jaunes) contenant des descendants de PBRIC et deux parents PB260, RRIC100

Rack3 + Rack4 des échantillons de PBRIM et un parent RRM600

Rack1												Rack2												Rack3												Rack4											
1	9	17	25	33	41	49	57	65	73	81	89	97	105	113	121	129	137	145	153	161	169	177	185	1	9	17	25	33	41	49	57	65	73	81	89	97	105	113	121	129	137	145	153	161	169	177	185
2	10	18	26	34	42	50	58	66	74	82	90	98	106	114	122	130	138	146	154	162	170	178	186	2	10	18	26	34	42	50	58	66	74	82	90	98	106	114	122	130	138	146	154	162	170	178	186
3	11	19	27	35	43	51	59	67	75	83	91	99	107	115	123	131	139	147	155	163	171	179	187	3	11	19	27	35	43	51	59	67	75	83	91	99	107	115	123	131	139	147	155	163	171	179	187
4	PB	20	28	36	44	52	60	68	76	84	92	100	108	116	124	132	140	148	156	164	172	180	188	4	RIM	20	28	36	44	52	60	68	76	84	92	100	108	116	124	132	140	148	156	164	172	180	188
5	RIC	21	29	37	45	53	61	69	77	85	93	101	109	117	125	133	141	149	157	165	173	181	189	5	13	21	29	37	45	53	61	69	77	85	93	101	109	117	125	133	141	149	157	165	173	181	189
6	14	22	30	38	46	54	62	70	78	86	94	102	110	118	126	134	142	150	158	166	174	182	190	6	14	22	30	38	46	54	62	70	78	86	94	102	110	118	126	134	142	150	158	166	174	182	190
7	15	23	31	39	47	55	63	71	79	87	95	103	111	119	127	135	143	151	159	167	175	183	191	7	15	23	31	39	47	55	63	71	79	87	95	103	111	119	127	135	143	151	159	167	175	183	191
8	16	24	32	40	48	56	64	72	80	88	96	104	112	120	128	136	144	152	160	168	176	184	192	8	16	24	32	40	48	56	64	72	80	88	96	104	112	120	128	136	144	152	160	168	176	184	192

Distribution  
au robot

Plaque384

1	1	9	9	17	17	25	25	33	33	41	41	49	49	57	57	65	65	73	73	81	81	89	89
97	97	105	105	113	113	121	121	129	129	137	137	145	145	153	153	161	161	169	169	177	177	185	185
2	2	10	10	18	18	26	26	34	34	42	42	50	50	58	58	66	66	74	74	82	82	90	90
98	98	106	106	114	114	122	122	130	130	138	138	146	146	154	154	162	162	170	170	178	178	186	186
3	3	11	11	19	19	27	27	35	35	43	43	51	51	59	59	67	67	75	75	83	83	91	91
99	99	107	107	115	115	123	123	131	131	139	139	147	147	155	155	163	163	171	171	179	179	187	187
4	4	PB	RIM	20	20	28	28	36	36	44	44	52	52	60	60	68	68	76	76	84	84	92	92
100	100	108	108	116	116	124	124	132	132	140	140	148	148	156	156	164	164	172	172	180	180	188	188
5	5	RIC	13	21	21	29	29	37	37	45	45	53	53	61	61	69	69	77	77	85	85	93	93
101	101	109	109	117	117	125	125	133	133	141	141	149	149	157	157	165	165	173	173	181	181	189	189
6	6	14	14	22	22	30	30	38	38	46	46	54	54	62	62	70	70	78	78	86	86	94	94
102	102	110	110	118	118	126	126	134	134	142	142	150	150	158	158	166	166	174	174	182	182	190	190
7	7	15	15	23	23	31	31	39	39	47	47	55	55	63	63	71	71	79	79	87	87	95	95
103	103	111	111	119	119	127	127	135	135	143	143	151	151	159	159	167	167	175	175	183	183	191	191
8	8	16	16	24	24	32	32	40	40	48	48	56	56	64	64	72	72	80	80	88	88	96	96
104	104	112	112	120	120	128	128	136	136	144	144	152	152	160	160	168	168	176	176	184	184	192	192

## 5. Références

- Blanc, G., Rodier-Goud, M., Y.J, L., Clément-Demange, A., and Seguin, M. (2001). Application des microsatellites à l'étude du brassage génétique dans un verger à graines d'hévéa. In: Rapport biennal d'activités scientifiques du Cirad-cp en 1998-1999.
- Beckmann, J. S., and Soller, M. (2003). Restriction fragment length polymorphisms in genetic improvement : methodologies, mapping and costs. *Theor. Appl. Genet.* 67:35-43.
- Churchill, G. A., and Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics* 138:963-971.
- Clément-Demange, A., Priyadarshan, P., M, Tran, T., Thuy, Hoa, and Venkatachalam, P. (2007). Hevea rubber breeding and genetics. *Plant Breeding Reviews, Edited by Jules Janick, ISBN 978-0-470-05241-9, John Wiley and Sons, Inc, Volume 29, 177-283.*
- de Vienne, D. E. (2003). Molecular markers in plant genetics and biotechnology. Inra Editions. ISBN: 2-7380-0776-7, 200 pp.
- de Vienne, D. E. (1995). Les marqueurs moléculaires en génétique et biotechnologies végétales. Inra Editions. ISBN: 2-7380-0776-7, 200 pp.
- <http://faostat.fao.org/>
- Geldermann, H. (1975). Investigations on inheritance of quantitative characters in animal by gene markers. I. Methods. *Theor. Appl. Genet.* 46:319-330.
- Grattapaglia, D., and Sederoff, R. (1994). Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross mapping strategy and RAPD markers. *Genetics* 137:1121-1137.
- Haldane, J. B. S. (1919). The combination of linkage values, and the calculation of distance between linked factors. *J. Genet.* 8, 299-309.
- Henderson, C. R. (1984). Applications of Linear Models in Animal Breeding. University of Guelph.
- Kosambi, D. D. (1944). The estimation of map distance from recombination values. *Ann. Eugen.* 12, 172-175.
- Lander, E. S., Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121, 185-199.
- Leconte, A., Lebrun, P., Nicolas, D., and Seguin, M. (1994). Electrophoresis : application to hevea clone identification. *Plantations, Recherche, Developpement* v. 1(2) p. 28-36.

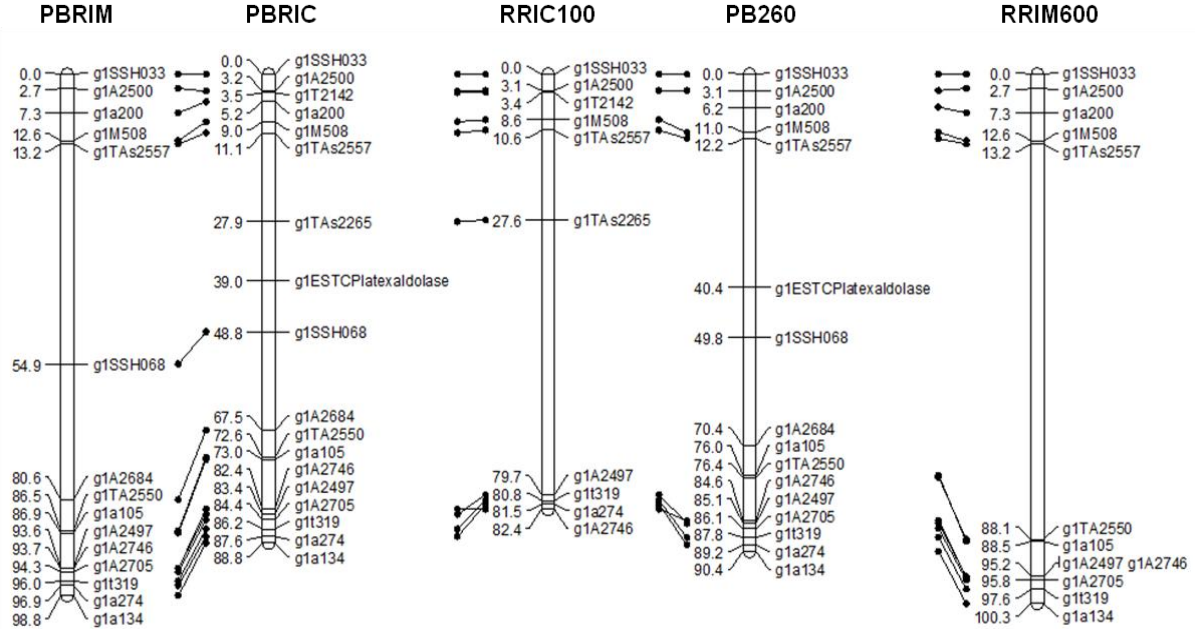
- Le Guen, V., Garcia, D., Doaré, F., Mattos, C. R. R., Condina, V., Couturier, C., Chambon, A., Weber, C., Espeout, S., and Seguin, M. (2011). A rubber tree's durable resistance to *Microcyclus ulei* is conferred by a qualitative gene and a major quantitative resistance factor. *Tree, Genetics & Genomes*, DOI 10.1007/s11295-011-0381-7.
- Le Guen, V., Gay, C., Xiong, T. C., Souza, L. M., Rodier-Goud, M., and Seguin, M. (2010). Development and characterization of 296 new polymorphic microsatellite markers for rubber tree (*Hevea brasiliensis*). *Plant Breeding* doi:10.1111/j.1439-0523.2010.01774.x.
- Le Guen, V., Doaré, F., Weber, C., and Seguin, M. (2009). Genetic structure of Amazonian populations of *Hevea brasiliensis* assessed by SSR markers, and application to germplasm management. *Tree Genetics and Genomes*, 5(4), 673-683.
- Le Guen, V., Garcia, D., Mattos, C., R, R, Doare, F., Lespinasse, D., and Seguin, M. (2007). Bypassing of a polygenic *Microcyclus ulei* resistance in rubber tree, analyzed by QTL detection. *New Phytologist* **173**, 335-345.
- Le Guen, V., Lespinasse, D., Oliver, G., Rodier Goud, M., Pinard, F., and Seguin, M. (2003). Molecular mapping of genes conferring field resistance to South American Leaf Blight (*Microcyclus ulei*) in rubber tree. *Theoretical and Applied Genetics* **108**, 160-167.
- Lekawipat N., Teerawatanasuk K., Rodier-Goud M., Seguin M., Vanachit A., Toojinda T. et Tragoonrung S., 2003. Genetic diversity analysis of wild germplasm and cultivated clones of *Hevea brasiliensis* Müll. Arg. by using microsatellite markers. *J. Rubb. Res.*, 6 (1) : 36-47.
- Lespinasse, D., Grivet, L., Troispoux, V., Rodier Goud, M., Pinard, F., and Seguin, M. (2000a). Identification of QTLs involved in the resistance to South American leaf blight (*Microcyclus ulei*) in the rubber tree. In "Theoretical and applied genetics.", Vol. Apr 2000. v. 100 (6) p. 975-984.
- Lespinasse, D., Rodier, G. M., Grivet, L., Leconte, A., Legnate, H., and Seguin, M. a. (2000b). A saturated genetic linkage map of rubber tree (*Hevea* spp.) based on RFLP, AFLP, microsatellite, and isozyme markers. *Theoretical and Applied Genetics* **100**, 127-138.
- Nguyen Thi Hue (1997). Hévée – Manuel général et technique. Edition éducative vietnamienne. Sax, K. (1923). The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8 : 552-560.
- Rattanawong, R. (2012). QTL mapping in *Hevea brasiliensis* for analysing the genetic determinism of growth, latex production, and the macromolecular structure of natural

- rubber. Kasetsart University - A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy, Graduate School, International Program, Department of Horticulture, Faculty of Tropical Agriculture, 4<sup>th</sup> April 2012.
- RRIV (2010). Recherche sur la sélection de variétés d'hévéa atteignant un rendement de latex de 3-3,5 tonnes/ha/an. Programme du Ministère de l'Agriculture et du Développement rural (Vietnam) sur la période 2006 – 2010. Rapport 2010.
- Schultes, R. E. (1990). A Brief Taxonomic View of the Genus *Hevea*. Kuala Lumpur, MRRDB, MRRDB Monograph n 14, 60 p. ISSN 0127-7782.
- Seguin, M., Flori, A., Legnaté, H., and Clément-Demange, A. (1999). L'hévéa. In: "Diversité génétique des plantes tropicales cultivées. Hamon, Seguin, Perrier et Glaszmann Eds", Cirad, Ird, Collection "Repères", pp.241-269.
- Seguin, M., Rodier-Goud, M., and Lespinasse, D. (1997). SSR markers in rubber tree (*Hevea brasiliensis*): facilitated and enhanced mapping using heteroduplex formation and DNA template mixing. Plant and Animal Genome V. Conference, San Diego, CA, USA, January 12-16, 1997, Poster n° 61. Abstract p. 66.
- Tran Dinh Minh (2012). Cartographie génétique de la famille PB260 x RRIC100 (*Hevea brasiliensis*) par marqueurs microsatellites et électrophorèse capillaire. Rapport de Master 1, Université Montpellier 2.
- Tran Thi Thuy Hoa (1998). Recherche sur l'amélioration de la pollinisation manuelle de l'hévéa au Vietnam. Thèse de doctorat, Université d'Agro-foresterie de Hochiminh-ville.
- Van Ooijen, J. W. (1992). Accuracy of mapping quantitative trait loci in autogamous species. Theor. Appl. Genet (1992), 84 : 803-811.
- Van Ooijen, J. W. (2004). MapQTL 5, Software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V., Wageningen, Netherlands.
- Varshney, R., Graner, A., and Sorrells, M. (2005). Genic microsatellite markers in plants: features and applications. Trends in biotechnology, 23:48-55.

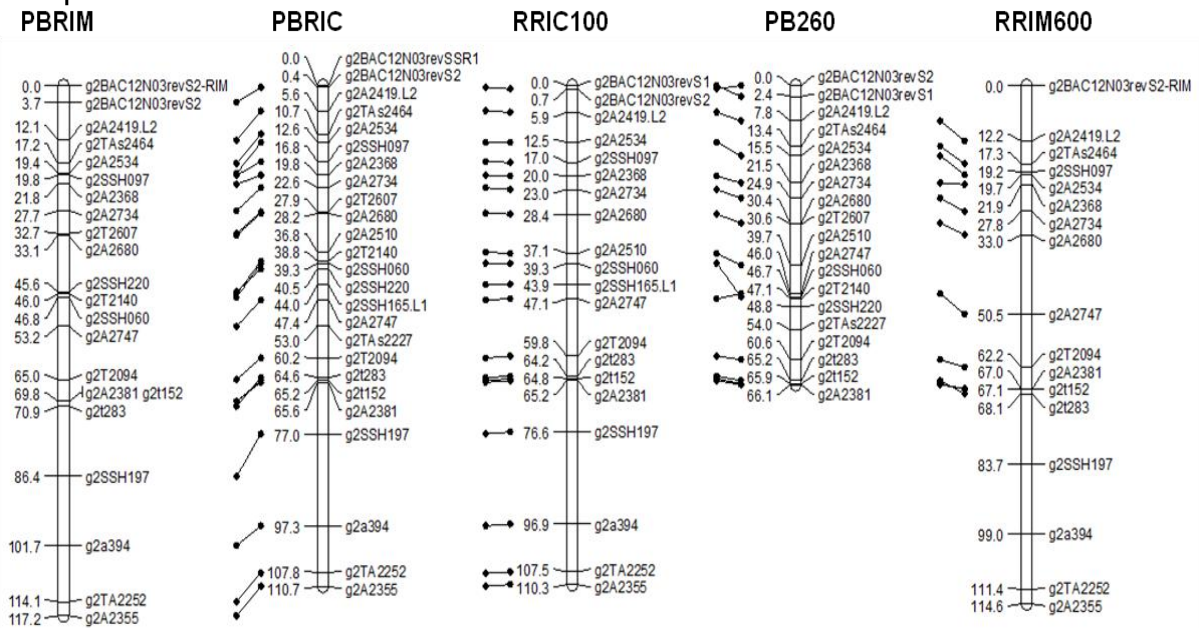
## 6. Annexes

### Annexe 1. Alignement des marqueurs sur les 18 groupes de liaison des cartes.

#### Groupe de liaison 1

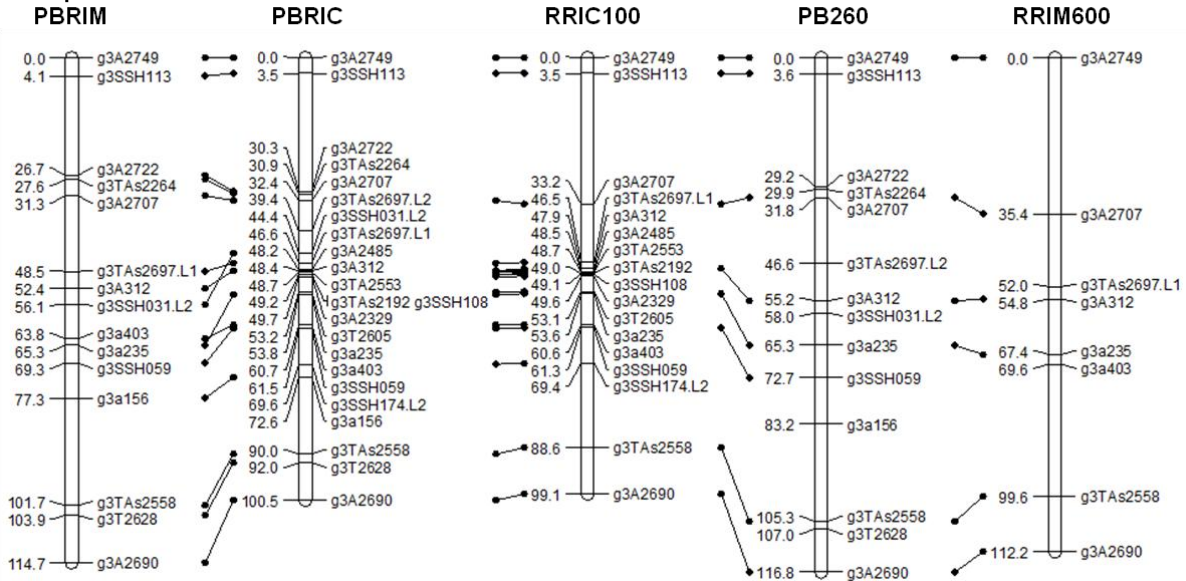


#### Groupe de liaison 2

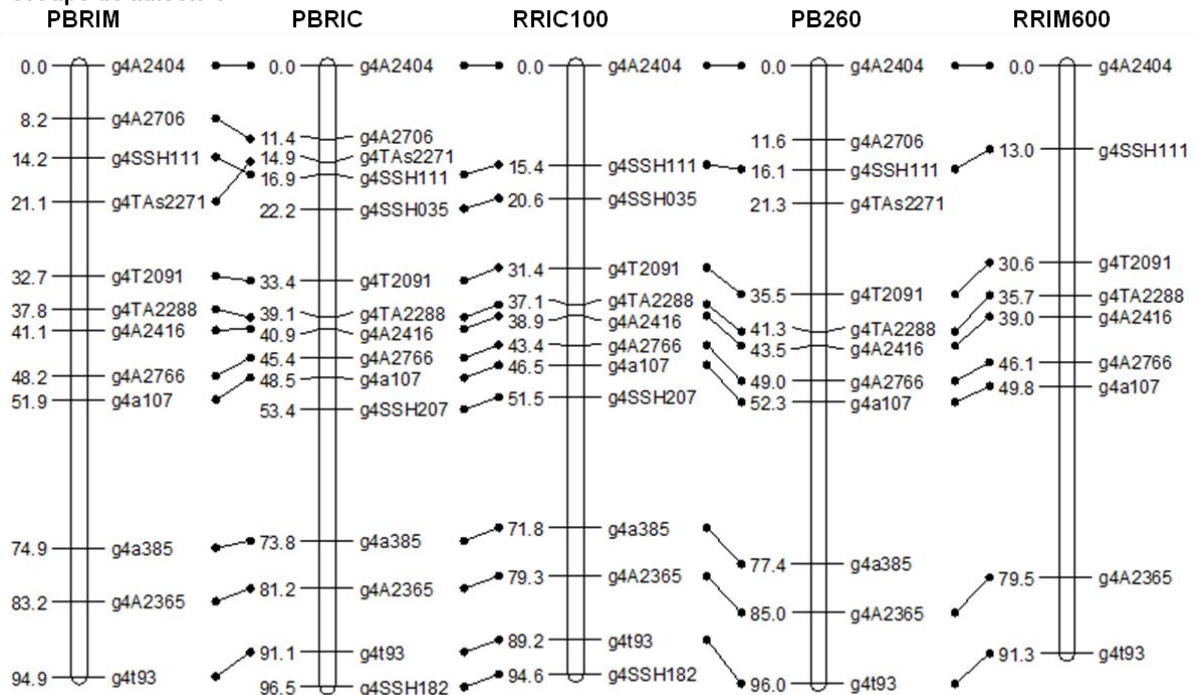




### Groupe de liaison 3

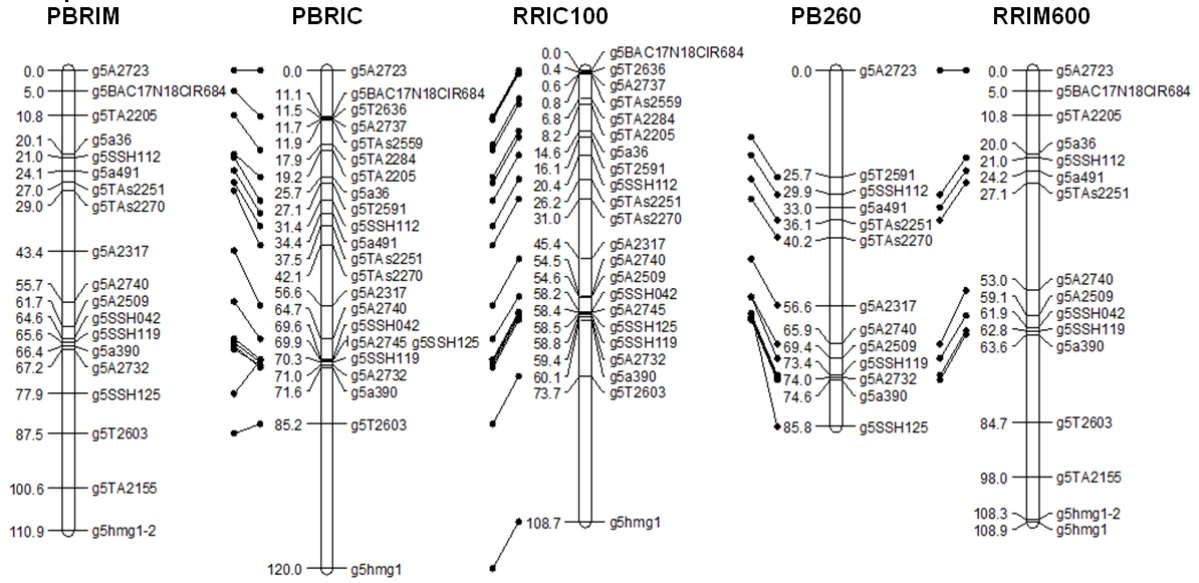


### Groupe de liaison 4

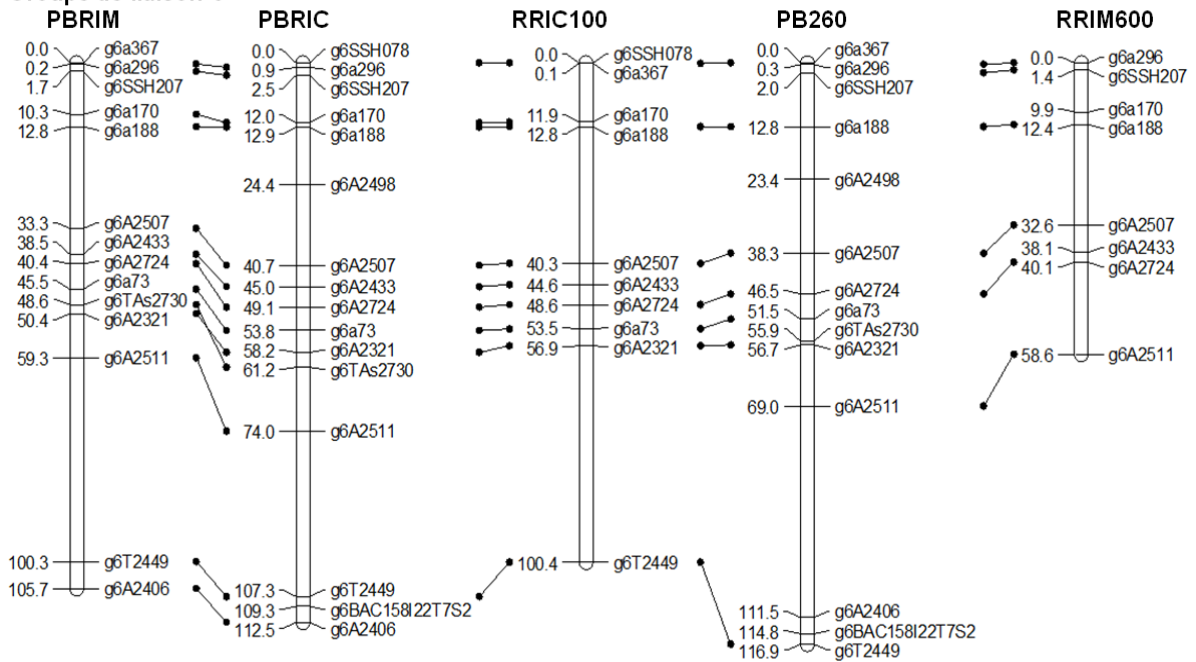




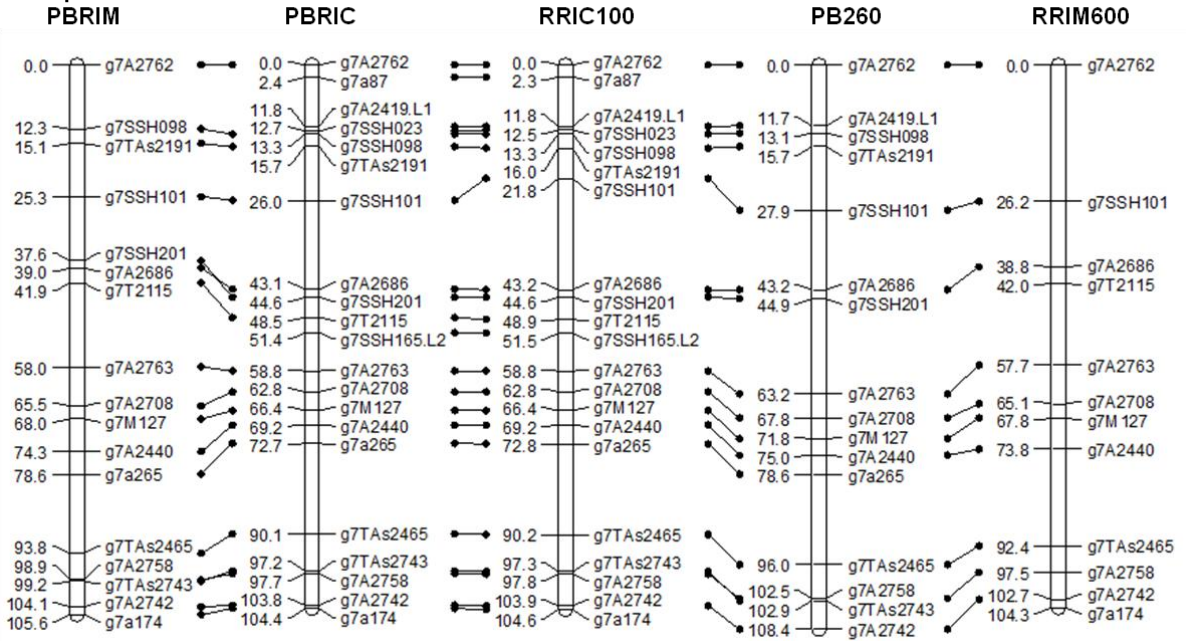
## Groupe de liaison 5



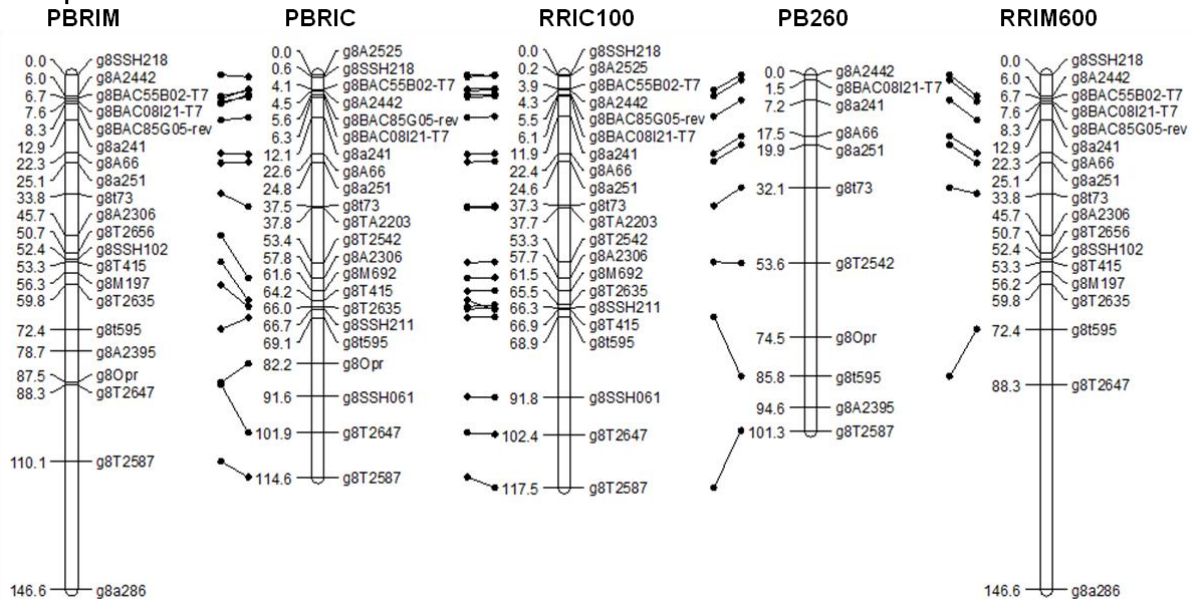
## Groupe de liaison 6



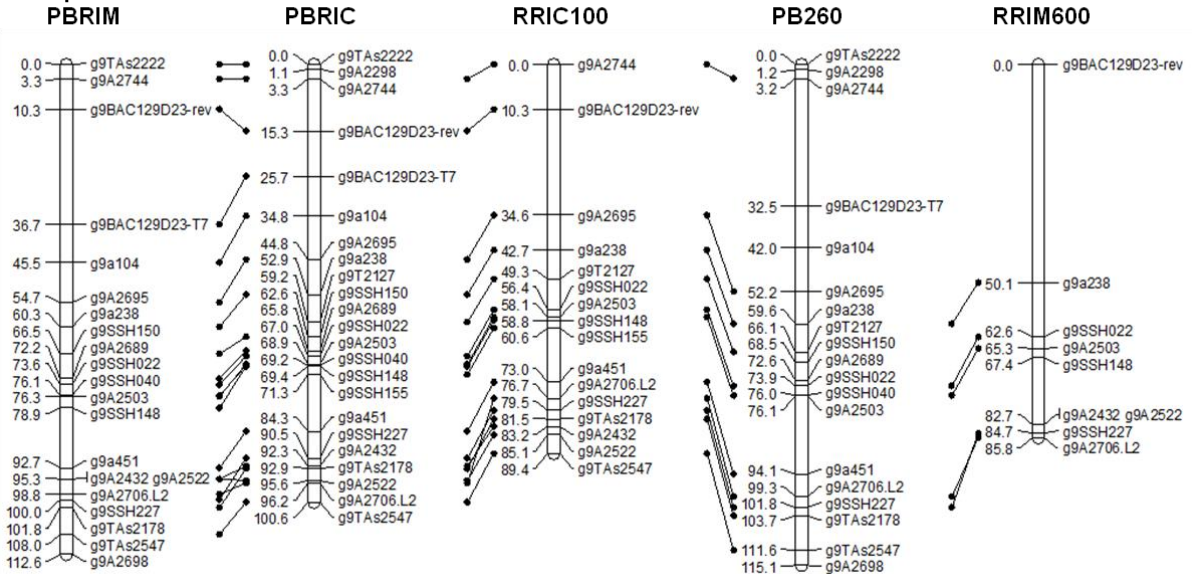
## Groupe de liaison 7



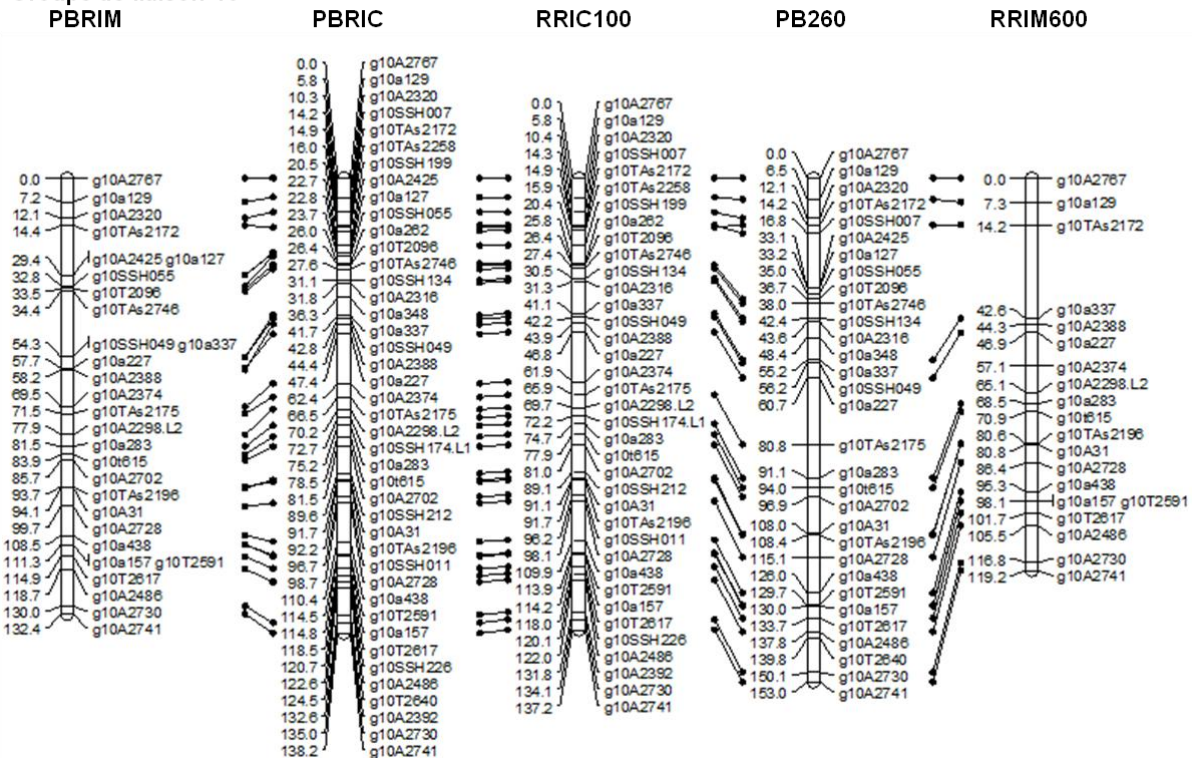
## Groupe de liaison 8



## Groupe de liaison 9

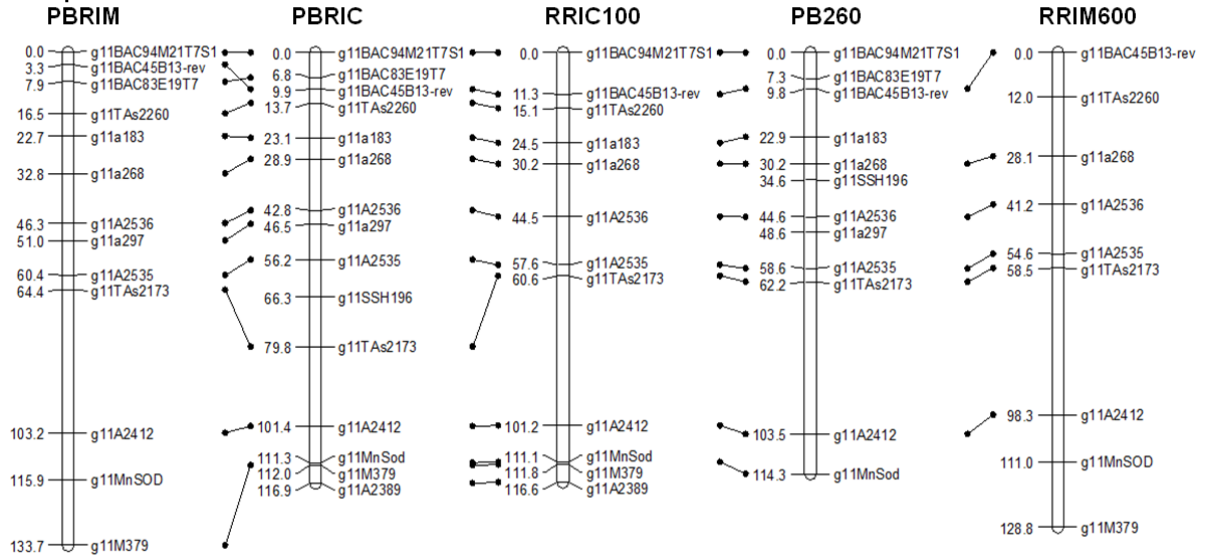


## Groupe de liaison 10

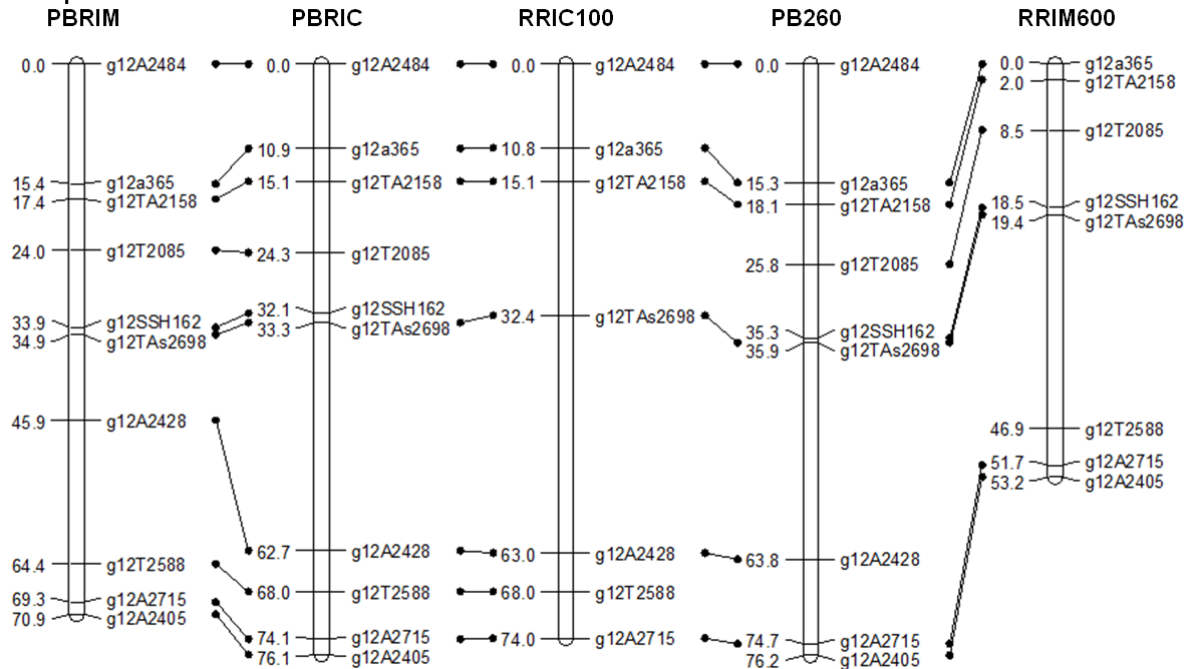




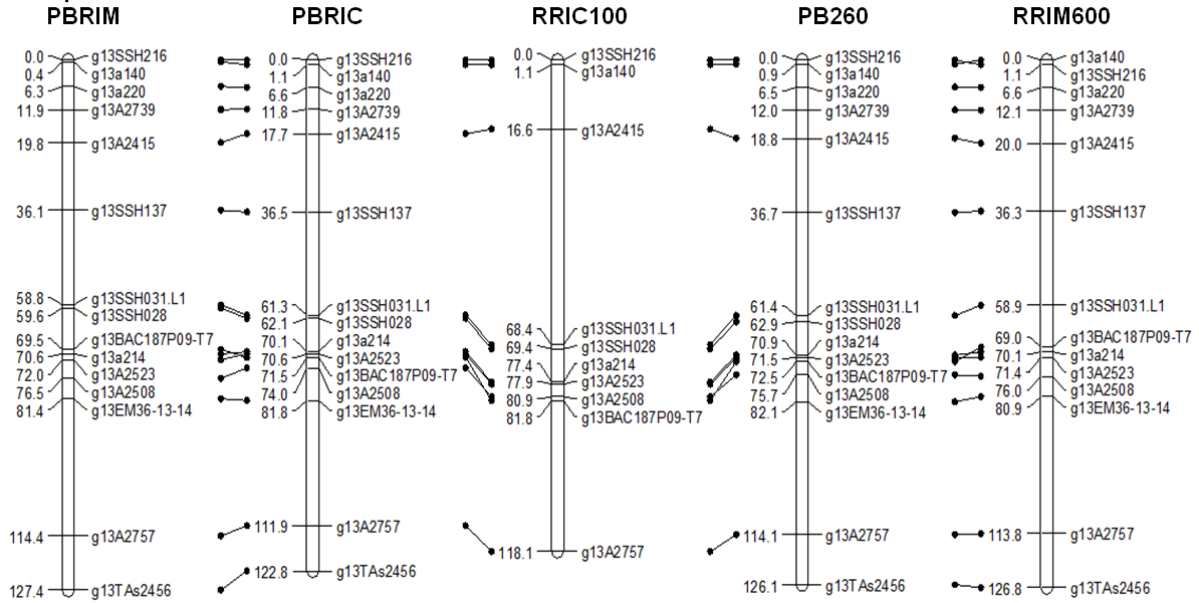
## Groupe de liaison 11



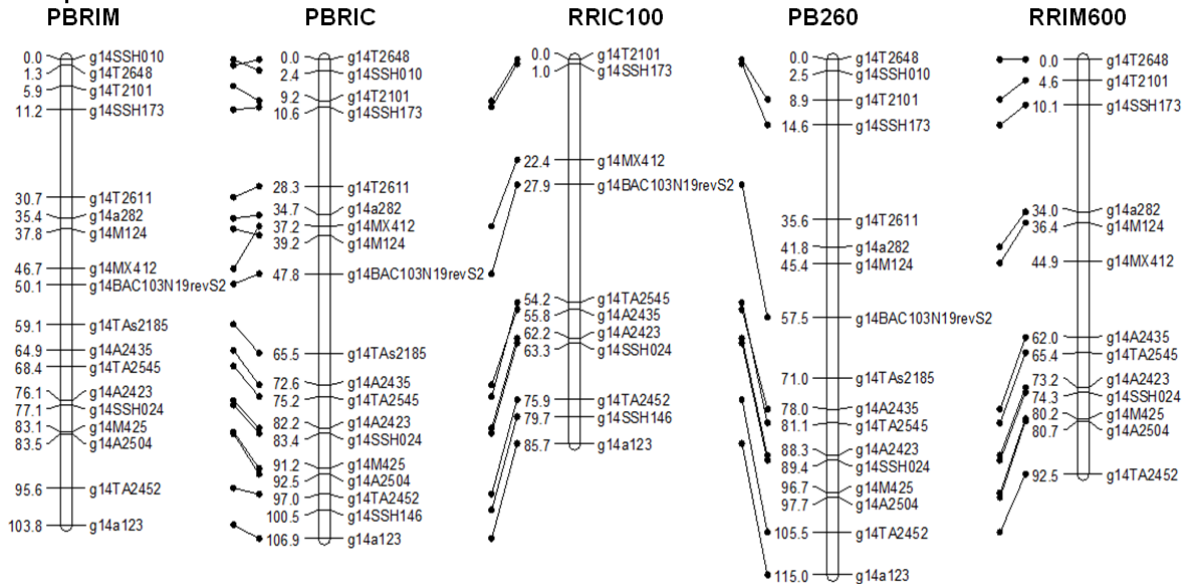
## Groupe de liaison 12



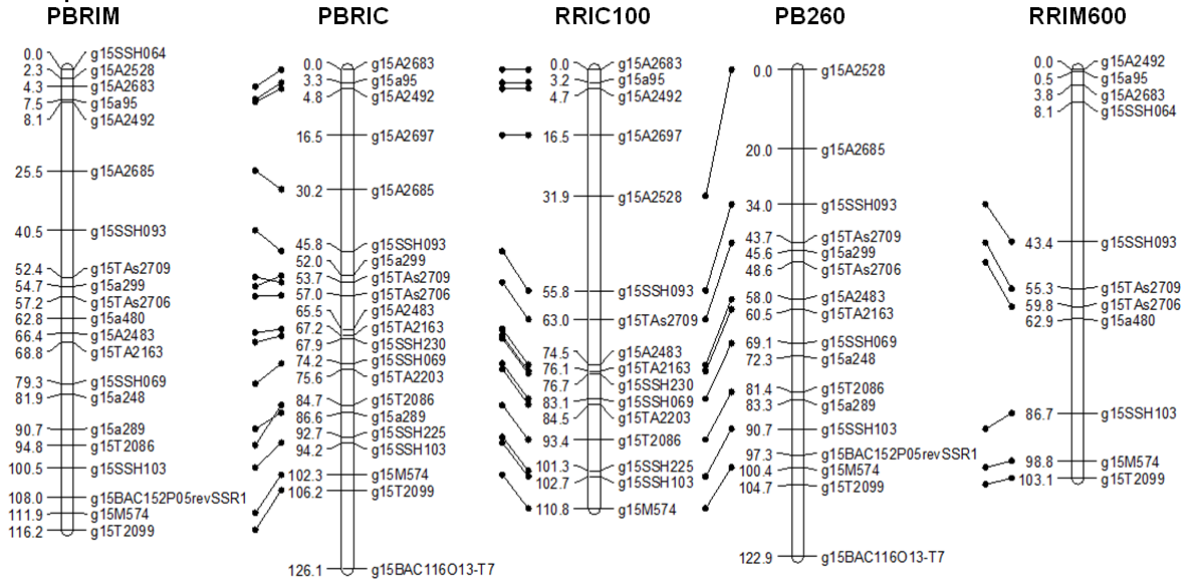
### Groupe de liaison 13



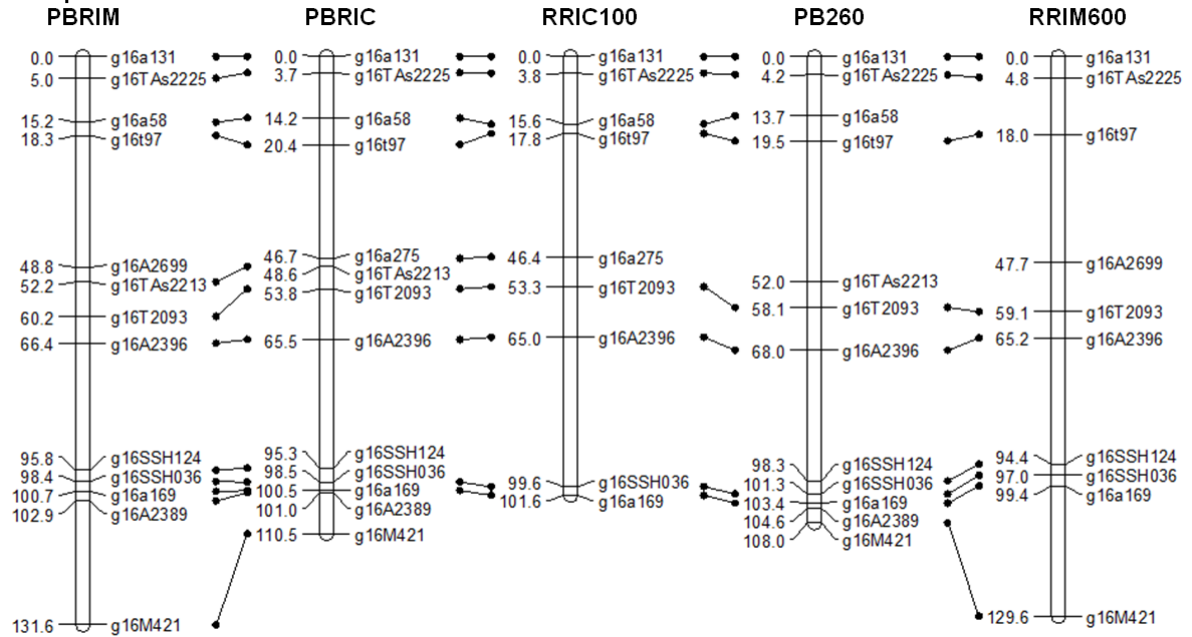
### Groupe de liaison 14



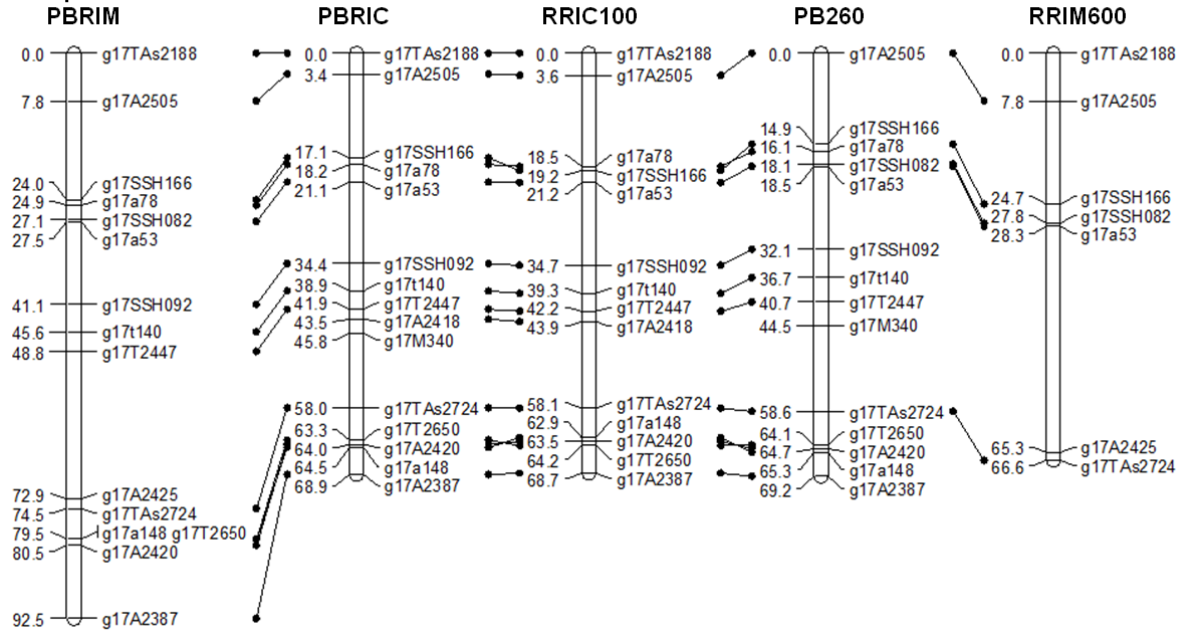
## Groupe de liaison 15



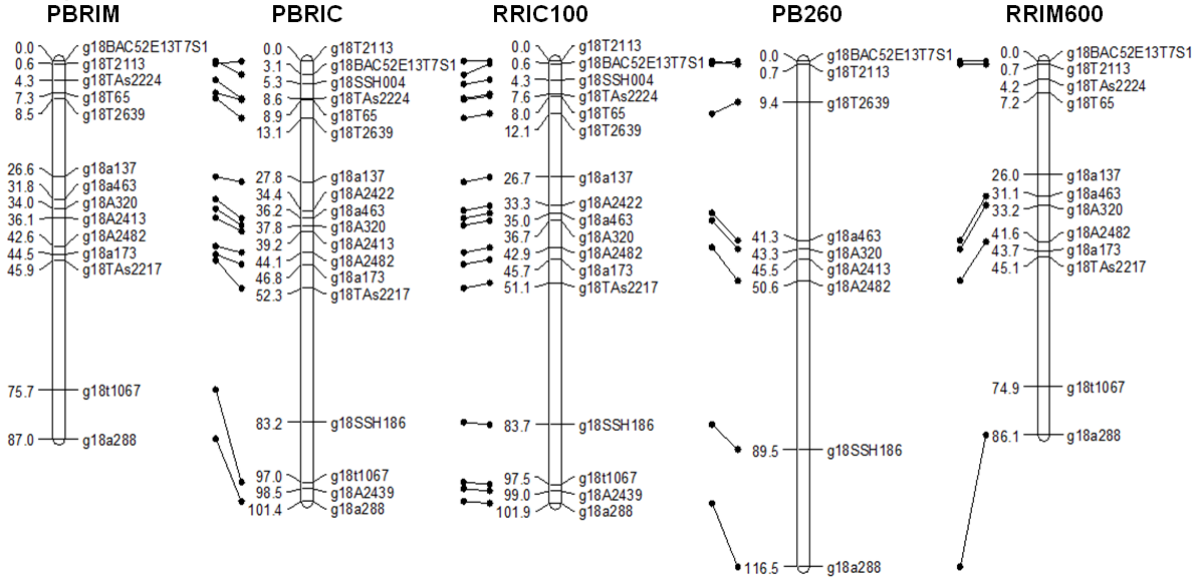
## Groupe de liaison 16



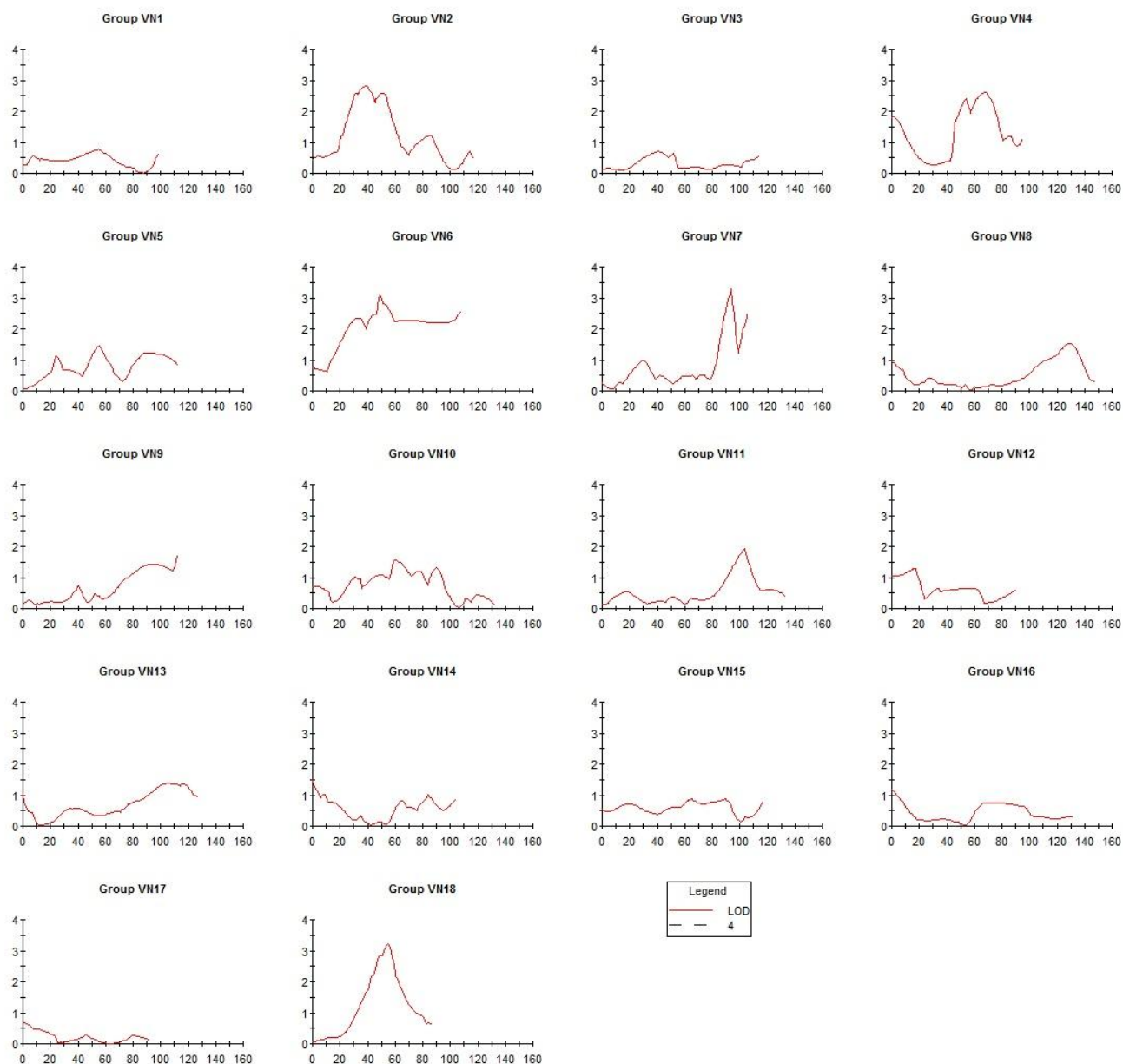
## Groupe de liaison 17



## Groupe de liaison 18

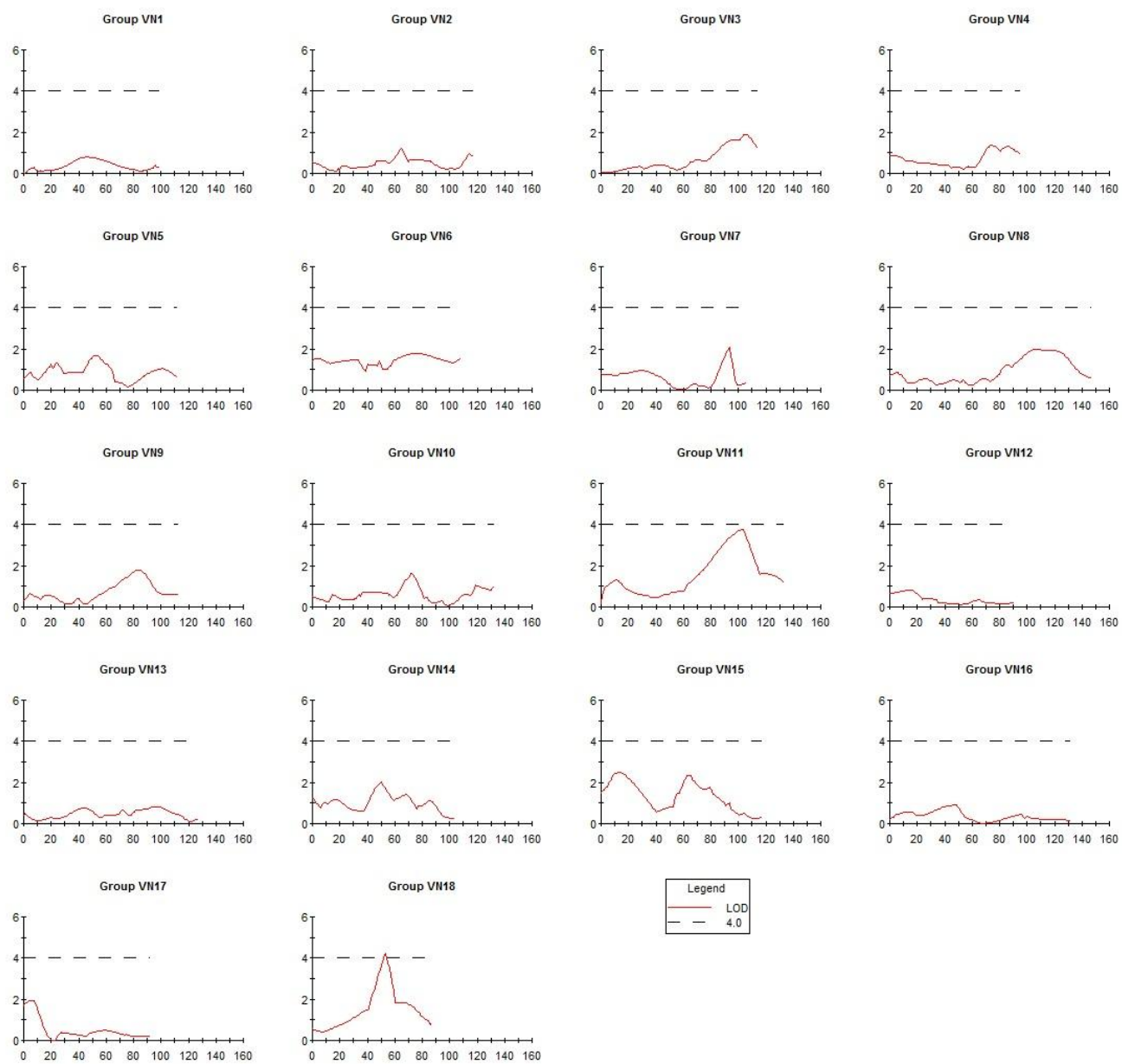


## Annexe 2. Détection de QTL de D4 pour la famille PBRIM (Côte d'Ivoire)





### Annexe 3. Détection de QTL de D9 pour la famille PBRIM (Côte d'Ivoire)



#### Annexe 4. Détection de QTL de D7 pour la famille PBRIC (Nigéria)

